

VISIT...

LANZAROTE
Caliente.COM

И. П. СТЕПАНЕНКО

ОСНОВЫ ТЕОРИИ ТРАНЗИСТОРОВ И ТРАНЗИСТОРНЫХ СХЕМ

*ИЗДАНИЕ ЧЕТВЕРТОЕ,
ПЕРЕРАБОТАННОЕ И ДОПОЛНЕННОЕ*



МОСКВА «ЭНЕРГИЯ»

1977

6Ф0.3

С79

УДК 621.382.3

Степаненко И. П.

С79 Основы теории транзисторов и транзисторных схем. Изд. 4-е, перераб. и доп. М., «Энергия», 1977.

672 с. с ил.

В книге проводятся анализ и расчет основных типов транзисторных усилителей, импульсных схем и источников питания. Анализ схем предшествует рассмотрению физических процессов в полупроводниковых диодах и транзисторах и характеристик диодов и транзисторов в качестве схемных элементов. Существенно переработана по сравнению с третьим изданием, вышедшим в 1973 г., первая часть книги, во вторую и третью части введены новые главы.

Книга предназначена для инженеров, аспирантов и студентов вузов, специализирующихся по микроэлектронике и прикладной электронике, вычислительной технике, автоматике и приборостроению.

С $\frac{30401-339}{051(01)-77}$ 186-77

6Ф0.3

В предисловии к 3-му изданию отмечалось, что данная книга помимо специализирующихся по дискретной транзисторной технике рассчитана также на один из новых типов специалистов, появившихся благодаря развитию микроэлектроники, — проектировщиков интегральных схем. При подготовке 4-го издания автор старался внести такие изменения, которые отвечали бы интересам и тех и других. А именно, в первой части («Транзисторы») существенно переработаны и дополнены разделы, посвященные эффекту поля, контакту «металл-полупроводник», а также § 5-4 и § 5-5 (в этих двух параграфах, в частности, учтен ряд особенностей, свойственных интегральному исполнению). Во всей первой части автор старался в максимальной степени отдавать предпочтение кремниевым приборам по сравнению с германиевыми.

Во вторую часть («Усилители») введены гл. 14 «Дифференциальный каскад», а также § 10-4 «Каскод». Эти дополнения давно назрели, так как оба каскада давно вошли в практику и дискретных, и интегральных схем. Особо хотелось бы отметить, что в книге впервые анализируются усилители постоянного тока на МДП транзисторах (§ 14-7). Глава 13 сокращена в объеме, поскольку материал о параллельно-балансных каскадах помещен в гл. 14, и по существу посвящена одноктактным усилителям постоянного тока.

В третьей части («Импульсные схемы») главное дополнение состоит в том, что в § 15-7 введен раздел «Токовые ключи», характеризующий ненасыщенный режим транзисторного ключа и одновременно дающий некоторое представление о логических схемах.

Как и в предыдущих изданиях, выполнен большой объем редакционной работы, включая сокращение текста в некоторых разделах и обновление литературы.

Автор отдает себе отчет в том, что структуру книги тоже желательно было бы усовершенствовать: материал, касающийся свойств полупроводников, превышает объемные нормы главы, материал по полевым транзисторам — нормы параграфов и т. п. Логиче-

ские элементы, ячейки памяти, компараторы и некоторые другие типы схем в книгу не включены сознательно: при отборе материала автор по-прежнему руководствуется тем, что книга посвящена основам транзисторной техники и не является энциклопедическим обзором всех известных вариантов приборов и схем. Кроме того, учитывается, что многие схемы перспективны лишь в интегральном исполнении, а микроэлектронике должна быть посвящена специальная книга.

В книге в целях сохранения преемственности с предыдущими изданиями оставлены прежние буквенные обозначения параметров полупроводниковых приборов (таблица соответствия обозначений, принятых в книге и по ГОСТ, приведена в приложении).

Автор будет признателен читателям за предложения и замечания по улучшению содержания книги, которые следует направлять по адресу: 113114, Москва, Шлюзовая наб., 10, издательство «Энергия».

Автор

Глава первая

ПОЛУПРОВОДНИКИ

1-1. ВВЕДЕНИЕ

Обычно к полупроводниковым материалам относят вещества, которые при комнатной температуре имеют удельное электрическое сопротивление ρ в пределах от 10^{-3} — 10^{-2} до 10^8 — 10^9 Ом·см *. Вещества со значительно меньшим сопротивлением (10^{-6} — 10^{-4} Ом·см) причисляют к проводникам (металлам), а со значительно большим (10^{10} — 10^{18} Ом·см) — к непроводникам (диэлектрикам). Металлы и диэлектрики являются «полюсами» наиболее общего класса веществ — полупроводников. Количество полупроводниковых материалов, известных в настоящее время, далеко превышает число металлов и диэлектриков. К полупроводникам относятся некоторые химические элементы (Si, Ge, Se), интерметаллические соединения (InSb, GaAs), окислы (Cu_2O , ZnO), сульфиды (CdS, ZnS), карбиды (SiC) и множество других химических соединений.

Различие между диэлектриками и полупроводниками в основном количественное. Различие же между металлами и полупроводниками отнюдь не сводится к различию удельных сопротивлений.

В отличие от чистых металлов сопротивление чистых полупроводников сильно зависит от температуры и, кроме того, с ростом температуры не увеличивается, а уменьшается. Так, для большинства металлов температурный коэффициент сопротивления составляет + (0,4—0,6) % на 1°C , а для полупроводников он может достигать — (5—6) % на 1°C и более.

При добавлении примеси в чистый металл удельное сопротивление образующегося сплава (нихрома, манганина и т. п.) больше удельного сопротивления каждого из компонентов, тогда как при добавлении примеси в чистый полупроводник его удельное сопротивление сильно уменьшается: например, $10^{-5}\%$ мышьяка в германии снижает сопротивление последнего в 200 раз.

Причины отмеченных особенностей полупроводников станут ясны после изучения их структуры и механизма проводимости.

* В полупроводниковой технике принято измерять удельное сопротивление для 1 см^3 материала ($1\text{ Ом}\cdot\text{см} = 10\text{ Ом}\cdot\text{мм}/\text{м}$).

В дальнейшем будут рассматриваться главным образом кремний и германий — основные материалы, используемые в современной транзисторной технике. Однако закономерности, характерные для этих материалов, распространяются в общем на весь класс *электронных полупроводников*, в которых электрический ток обусловлен только электронами, а не ионами [1, 2].

1-2. СТРУКТУРА ПОЛУПРОВОДНИКОВ И ТИПЫ ПРОВОДИМОСТИ

Полупроводники представляют собой кристаллы, которые можно разбить на множество повторяющихся однотипных *элементарных ячеек*¹. Такие ячейки — структурные «молекулы» кристалла — подобно химическим молекулам состоят из нескольких атомов [3, 4].

Кристаллическая решетка кремния называется тетраэдрической или решеткой типа алмаза. Она характерна также для германия и всех четырехвалентных элементов. В основе такой решетки лежит пространственная фигура — тетраэдр, в углах и центре которой расположены атомы (рис. 1-1, а). Характерная особенность тетраэдрической системы заключается в одинаковом расстоянии центрального атома от четырех угловых. Каждый угловой атом

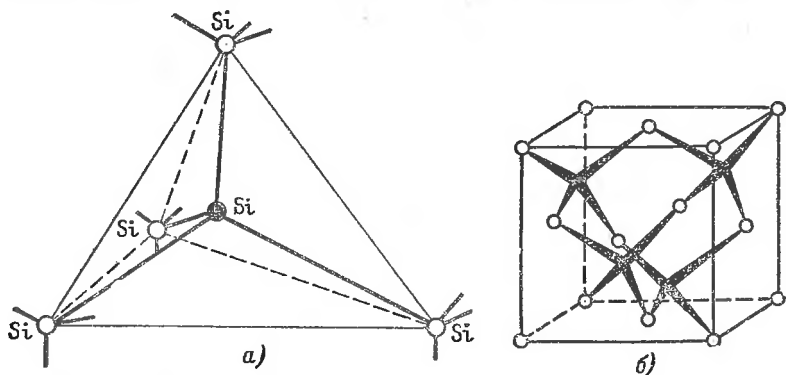


Рис. 1-1. Тетраэдрическая структура кристаллической решетки.

а — элементарный тетраэдр; б — элементарная ячейка.

в свою очередь служит центральным для других четырех ближайших атомов. Совокупность нескольких тетраэдров образует элементарную ячейку (рис. 1-1, б) кубической формы с размером ребра около 0,5 нм (*постоянная решетки*).

¹ Говоря о кристаллах, обычно подразумевают *монокристаллы*, которые в отличие от *поликристаллов* совершенно однородны, т. е. обладают идеальной, регулярной решеткой. Поликристаллы, несмотря на внешнюю однородность, имеют зернистую структуру, т. е. состоят из множества микрокристаллов, разделенных межкуристаллическими границами, равносильными дефектам решетки. Для изготовления обычных полупроводниковых приборов (диодов, транзисторов) поликристаллы непригодны.

Всем кристаллам, в том числе типичным полупроводникам, свойственна *анизотропия* — зависимость свойств от направления, в котором эти свойства рассматриваются. Анизотропия является естественным следствием того, что расстояния между смежными атомами различны в разных направлениях. Плоскости, лежащие внутри кристалла, и его наружные грани принято обозначать индексами Миллера [4] — совокупностью трех цифр, характеризующих ориентацию плоскости относительно атомов элементарной ячейки. Так, применительно к кубической решетке индекс (100) соответствует плоскости, проходящей через

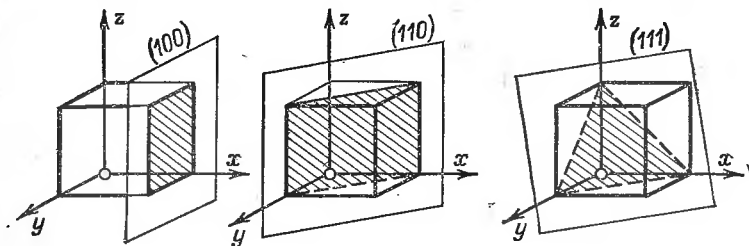


Рис. 1-2. Характеристические плоскости кубической решетки.

грань куба; индекс (110) — плоскости, проходящей через диагонали противоположных граней; а индекс (111) — плоскости, проходящей через три вершины, не связанные общими ребрами (рис. 1-2)¹. Направления, перпендикулярные указанным плоскостям, обозначаются теми же индексами, но в квадратных скобках. Наиболее «густым» по числу атомов на единицу длины является направление [111], наиболее «разреженным» — направление [100].

В последующих выводах мы будем пользоваться плоским эквивалентом тетраэдрической структуры (рис. 1-3), в котором сохранена главная особенность решетки типа алмаза — одинаковые расстояния между смежными атомами.

Связь атомов в рассматриваемой решетке устанавливается вследствие наличия специфических обменных сил, возникающих в результате попарного объединения валентных электронов. Эти силы отражены на рис. 1-3 сходящимися стрелками. Объединение электронов следует понимать так: пара электронов, объединенная на рис. 1-3 пунктиром, в равной степени принадлежит обоим соответствующим атомам или, образно говоря, поочередно примыкает то к одному, то к другому. Соответственно «в среднем» каждый атом обладает устойчивой восьмизлектронной оболочкой. Такая связь атомов, при которой каждый из них остается нейтральным, называется парноэлектронной, *ковалентной* [2] или просто валентной в отличие от *ионной* связи, обусловленной электрическими силами и имеющей место, например, в решетке каменной соли Na^+Cl^- , где атомы ионизированы.

¹ В общем виде индексы Миллера определяются следующим образом. Начало прямоугольной системы координат помещают в одном из узлов решетки; отрезки l , m , n , отсекаемые данной плоскостью на осях координат, измеряют в единицах постоянной решетки, обратные величины l^{-1} , m^{-1} , n^{-1} приводят к общему наименьшему знаменателю, после чего знаменатель отбрасывают. Числители образуют индексы Миллера.

Совершенно однородная структура (рис. 1-3) бывает у кристалла лишь при температуре абсолютного нуля. По мере нагревания полупроводника часть валентных связей нарушается под действием тепловых колебаний атомов в решетке. В корпускулярной интерпретации носителями энергии механических колебаний атомов являются *фононы* — акустические аналоги световых фотонов¹. Поэтому можно сказать, что при нагреве кристалла валентные связи нарушаются фононами, число и энергия которых растут с температурой.

Нарушение валентных связей приводит к одновременному образованию свободных электронов и пустых мест — *дырок* вблизи

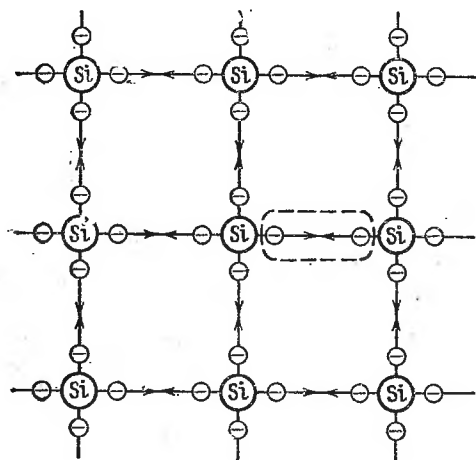


Рис. 1-3. «Плоский» эквивалент тетраэдрической решетки с валентными связями атомов.

тех атомов, от которых оторвались электроны (рис. 1-4). Такая дырка ведет себя подобно частице с элементарным положительным зарядом. Она, так же как и электрон, совершает хаотическое движение в течение некоторого времени после своего появления (*времени жизни*), а затем рекомбинирует с одним из свободных электронов. На рис. 1-5 показана схема движения дырки в решетке кремния как результат последовательного замещения ее электронами, принадлежащими разным атомам: исходная дырка в атоме *I* замещается од-

ним из электронов атома *II* и «переходит» к атому *II*; затем дырка в атоме *II* замещается одним из электронов атома *VI* и «переходит» к атому *VI* и т. д. Создается впечатление, что одна и та же дырка непрерывно движется по пути *I—II—VI—X—VII—III*, тогда как на самом деле совершаются дискретные переходы электронов в обратном направлении: *II—I*, *VI—II*, *X—VI* и т. д.

Таким образом, в полупроводниках имеются два типа подвижных носителей заряда — электроны и дырки.

При нагревании абсолютно чистого и однородного полупроводника (*собственный* полупроводник) свободные электроны и дырки всегда образуются парами, как следует из рис. 1-4. Число этих

¹ Тепловые колебания решетки и соответственно фононы делятся на два типа: *акустические* (низкочастотные) и *оптические* (высокочастотные). Первые являются результатом синфазных, а вторые — противофазных колебаний атомов в элементарной ячейке.

пар в стационарном режиме определяется равновесием между процессами *термогенерации* и *рекомбинации* носителей. Проводимость собственного полупроводника, обусловленную парными носителями теплового происхождения, называют *собственной*.

Проводимость, обусловленную наличием примесных атомов, нарушающих структуру кристаллической решетки, называют *примесной*. Заметим, что слово «примесь» не всегда следует понимать буквально. В ряде случаев такие же последствия, как наличие примесных атомов, могут вызывать различные *дефекты* решетки: избыток одного из основных компонентов вещества (например, кислорода в закиси меди), смещение некоторых узлов решетки и др. Поэтому более точен общий термин — *дефектная* проводимость.

Атомы примеси могут по-разному располагаться в решетке собственного полупроводника. В кремнии и германии примесные атомы обычно замещают часть основных атомов в узлах решетки. Результаты такого замещения зависят от типа примеси.

Если ввести в кремний атом пятивалентного фосфора, то четыре из его валентных электронов вступят в связь с четырьмя электронами соседних атомов кремния (рис. 1-6, а) и образуют устойчивую оболочку из восьми электронов. Девятый электрон в этой комбинации оказывается слабо связанным с ядром, легко отрывается фононами и делается свободным. При этом примесный атом превращается в неподвижный ион с единичным положительным зарядом.

Свободные электроны примесного происхождения добавляются к собственным свободным электронам, порожденным термогенерацией, поэтому проводимость полупроводника делается преимущественно электронной. Такие полупроводники называются *электронными*¹ или *типа n*. Примеси, обуславливающие электронную проводимость, называются *донорными* («отдающими» электроны). Донорами по отношению к кремнию и германию, помимо фосфора, могут быть сурьма, мышьяк и некоторые другие элементы.

Однако если ввести в кремний атом трехвалентного бора, результат будет иным. Для валентной связи бора с четырьмя соседними атомами кремния требуется по-прежнему образование устойчивой восьмиэлектронной оболочки, т. е. нужен дополнительный

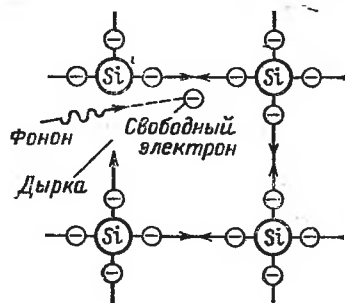


Рис. 1-4. Процесс образования пары электрон — дырка в решетке под действием фонона (или фотона).

¹ Этот термин не следует путать с более общим термином «электронный полупроводник», подразумевающим отсутствие движения ионов (см. § 1-1).

электрон. Этот электрон отбирается из основной решетки (рис. 1-6, б) и, будучи связанным, превращает атом бора в неподвижный отрицательный ион. На том месте, откуда пришел электрон, образуется свободная дырка, которая добавляется к собственным дыркам, порожденным термогенерацией. Такие полупроводники (с преимущественно дырочной проводимостью) называются *дырочными* или *типа p*, а соответствующие примеси — *акцепторными* («принимающими» электроны). Помимо бора акцепторами по отношению к кремнию и германию служат алюминий, галлий, индий и некоторые другие элементы.

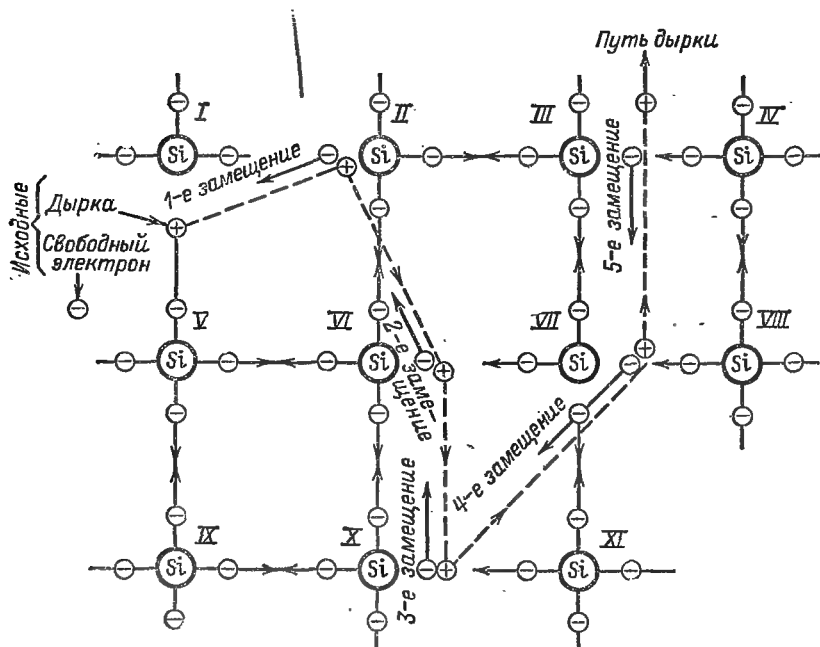


Рис. 1-5. Схема движения свободной дырки в кристаллической решетке.

Отрыв «лишнего» электрона от донора и «недостающего» электрона для акцептора требует затраты некоторой энергии (энергия ионизации или активации примеси). Эта энергия зависит от типа и концентрации примеси, а также от диэлектрической проницаемости основного материала, которая влияет на силу электрического притяжения между электроном и атомным «остатком» [5]. При нулевой температуре ($T = 0$ К) ионизация не может иметь места; с ростом температуры все большая часть примесных атомов ионизируется фонами. В кремнии и германии при комнатной температуре примесные атомы III и V групп ионизированы практически полностью.

Полезно сделать несколько замечаний относительно полупроводников сложного состава, а именно двойных соединений типа $A^{III}B^V$, где A и B — обозначения двух разных атомов, индексы означают валентность. К числу таких соединений относятся арсенид галлия GaAs, антимонид индия InSb и др. Полупроводники со структурой $A^{III}B^V$ также имеют тетраэдрическую кристаллическую решетку типа алмаза, но атомы A и B равномерно распределены в ней так, что в центре каждого А-тетраэдра находится атом B , а в центре каждого В-тетраэдра — атом A ; тетраэдры A и B переплетаются подобно звеньям цепи. Связь между атомами A и B ковалентна, так как суммарное число валентных электронов у пары атомов A и B равно восьми, а это соответствует устойчивой восьми-электронной оболочке. Донорными примесями для соединений $A^{III}B^V$ являются элементы VI группы (теллур), а акцепторными — элементы II группы (кадмий,

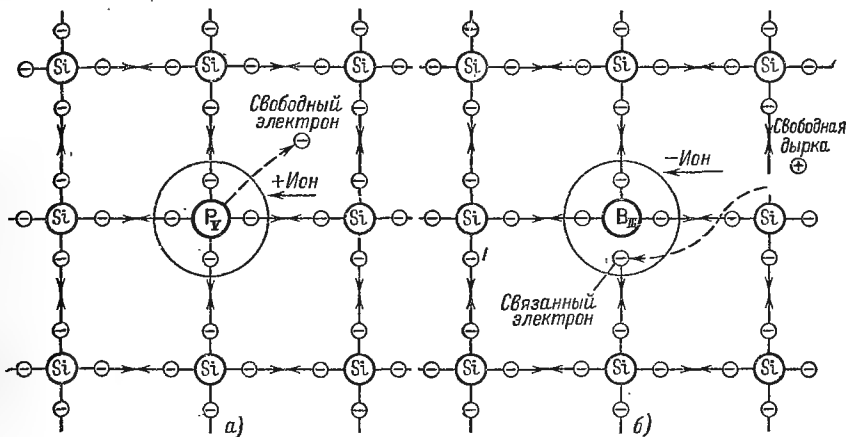


Рис. 1-6. Замещение примесными атомами основных атомов в решетке.

а — донорная примесь (образуются свободный электрон и неподвижный положительный ион); б — акцепторная примесь (образуются свободная дырка и неподвижный отрицательный ион).

цинк). Донорные атомы замещают атомы B , а акцепторные — атомы A с образованием соответственно свободного электрона или свободной дырки.

Поскольку в примесных полупроводниках один тип подвижных носителей заряда превалирует над другим, принято называть те носители, которые составляют большинство, *основными*, а те, которые составляют меньшинство, — *неосновными*. Так, в p -полупроводнике основные носители — дырки, неосновные — электроны.

1-3. ЭНЕРГЕТИЧЕСКИЕ ЗОНЫ ТВЕРДОГО ТЕЛА

Количественный анализ полупроводников и полупроводниковых приборов базируется на зонной теории твердого тела.

Как известно, изолированный атом характеризуется дискретным спектром энергий, разрешенных для электронов (рис. 1-7). Расстояния между последовательными уровнями непрерывно уменьшаются по мере увеличения энергии. «Потолком» энергетического спектра является уровень ионизации, на котором электрон делается свободным и может покинуть атом. Заполненные уровни

образуют электронные оболочки атома, которые обозначают цифрами 1, 2, 3... или буквами *K, L, M...* Оболочки, начиная со второй, разбиваются на подоболочки: *2s, 2p, 3s, 3p, 3d* и т. д. Количество оболочек и подоболочек, заполненных электронами, зависит от порядкового номера элемента [6].

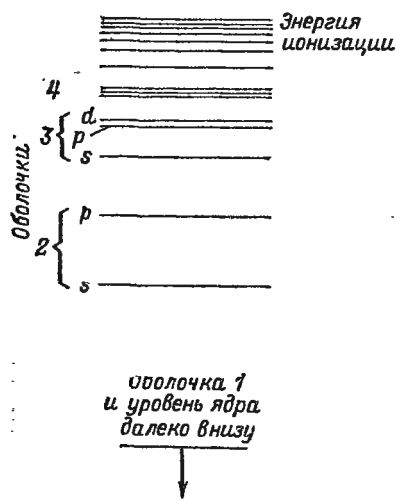


Рис. 1-7. Энергетические уровни изолированного атома.

В возбужденном состоянии атома его верхние уровни всегда свободны. Твердое тело представляет собой множество атомов, *сильно взаимодействующих* благодаря малым межатомным расстояниям. Поэтому всю совокупность атомов в куске твердого тела следует рассматривать как единое целое, как гигантскую молекулу, которая подобно атому характеризуется некоторым *единым* для всего тела энергетическим спектром. Особенность этого спектра в том, что он состоит не из дискретных разрешенных уровней, а из дискретных *разрешенных зон*. Каждая зона происходит от соответствующего атомного уровня, который как бы расщепляется при

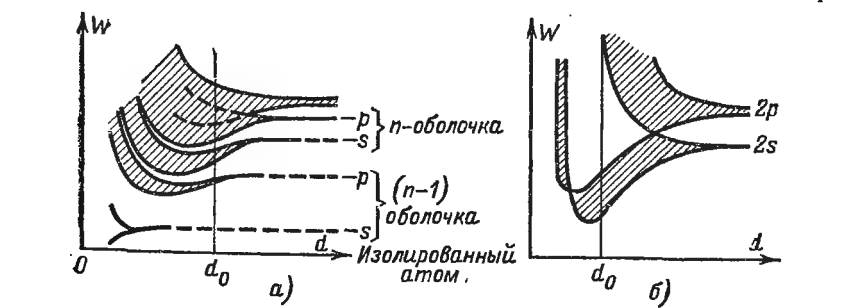


Рис. 1-8. Превращение разрешенных энергетических уровней отдельного атома в разрешенные энергетические зоны твердого тела.

а — общий случай; б — структура алмаза.

дуются с *запрещенными зонами* [2]. Ширина тех и других обычно не превышает нескольких электронвольт и не зависит от числа атомов в твердом теле, т. е. от его размеров.

Строго говоря, разрешенные зоны имеют дискретную структуру и состоят из столько уровней, сколько атомов имеется в дан-

ном теле [1]. Однако количество атомов даже в микроскопических объемах настолько велико, что энергетические расстояния между уровнями зоны в реальных случаях не превышают 10^{-17} эВ, т. е. разрешенные зоны практически можно считать сплошными.

Нижние энергетические уровни атомов обычно не образуют зон, так как внутренние электронные оболочки слабо взаимодействуют в твердом теле, будучи «экранированы» внешними оболочками. В связи с этим нижние уровни сохраняют свою «индивидуальность» и их показывают на зонной диаграмме в виде штриховой линии, где каждый штрих как бы соответствует одному атому (рис. 1-8).

В ряде случаев разрешенные зоны перекрываются, и тогда соответствующая запрещенная зона может отсутствовать. Такое перекрытие всегда имеет место в верхней части спектра, поскольку верхние уровни в отдельном атоме расположены весьма близко друг к другу. В результате энергетический спектр твердого тела содержит единую верхнюю зону и полное число зон в отличие от числа уровней в атоме оказывается конечным.

Энергетические «расстояния» между разрешенными зонами (т. е. ширина запрещенных зон) определяется энергией связи электронов с атомами решетки. Поэтому граничные энергетические уровни, образующие «дно» и «потолок» каждой разрешенной зоны, соответствуют чисто потенциальной энергии электронов, т. е. их неподвижному состоянию. Любой уровень, расположенный внутри разрешенной зоны, соответствует сумме потенциальной и кинетической энергии. Иначе говоря, кинетическая энергия электронов возрастает по мере удаления от границы в глубь зоны и достигает максимума в ее средней части.

Проводимость в твердом теле возможна лишь тогда, когда возможен переход электрона на ближайший энергетический уровень. Значит, в проводимости могут участвовать электроны только тех зон, в которых есть свободные уровни. Такие свободные уровни всегда имеются в верхней разрешенной зоне. Поэтому верхнюю зону твердого тела, не заполненную (или не полностью заполненную) электронами при нулевой абсолютной температуре, называют *зоной проводимости*.

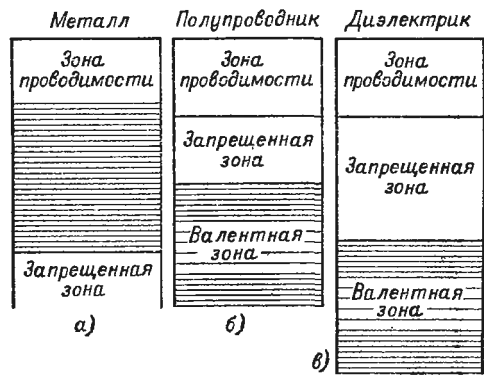


Рис. 1-9. Зонная структура при $T = 0$ К.
а — металла; б — полупроводника; в — диэлектрика.

Зону, ближайшую к зоне проводимости, называют *валентной*. При нулевой температуре она полностью заполнена. Следовательно, электроны этой зоны не могут участвовать в проводимости. Но при температуре, отличной от нуля, в верхней части валентной зоны образуются свободные уровни, и эта зона также может обусловить проводимость¹. Освобождение уровней в других, более глубоких зонах является особым случаем, на котором мы не будем останавливаться. Таким образом, все существенные процессы в полупроводниковых приборах можно изучить, рассматривая только две смежные зоны: зону проводимости и валентную.

Зонная структура твердого тела при нулевой температуре лежит в основе классификации металлов, полупроводников и диэлектриков (рис. 1-9).

У металлов зона проводимости и валентная зона взаимно перекрываются, поэтому даже при нулевой температуре в зоне проводимости находится значительное количество электронов и, следовательно, имеет место проводимость. У полупроводников и диэлектриков при нулевой температуре зона проводимости пуста и проводимость отсутствует, в этом их качественное отличие от металлов. Различия же между полупроводниками и диэлектриками в основном количественные и обусловлены значительно большей шириной запрещенной зоны у диэлектриков.

Поскольку зона проводимости практически сплошная, энергия электронов в этой зоне может меняться непрерывно, как у изолированных электронов в вакууме, поэтому электроны в зоне проводимости называются *свободными*. Термин «свободный» характеризует возможность перемещения электрона внутри твердого тела, но отнюдь не возможность вылета за пределы кристалла. Вылет электрона из твердого тела (*термоэмиссия*) возможен лишь в результате преодоления весьма высокого (несколько электрон-вольт) потенциального барьера на поверхности тела. При обычных рабочих температурах полупроводниковых приборов термоэмиссию можно не учитывать.

Таким образом, совокупность электронов в твердом теле можно уподобить *электронному газу*, который заключен в «сосуд», образованный внешними гранями кристалла. Наличие в этом «сосуде» множества неподвижных атомов — узлов решетки приводит к тому, что свойства электронов в твердом теле отличаются от их свойств в свободном пространстве — вакууме. Одно из важнейших отличий состоит в том, что масса электрона в кристалле, вообще говоря, не совпадает с его массой m в вакууме. Поэтому в теории твердого тела пользуются понятием *эффективной массы* m^* , которая зависит от ряда факторов (полная энергия, направление движения и др.).

¹ Заполненность валентной зоны соответствует в корпускулярном аспекте (§ 1-2) наличию у всех атомов устойчивых восьмизлектронных оболочек, обусловленных ковалентными связями, а освобождение уровней — разрыву этих связей и образованию дырок.

Различие между величинами m и m^* может быть не только количественным, но и качественным. А именно, эффективная масса электронов, энергия которых лежит в верхней части разрешенной зоны, оказывается отрицательной. Это значит, что такие электроны под действием ускоряющего электрического поля теряют кинетическую энергию, т. е. движутся в сторону потолка зоны, где энергия чисто потенциальная (см. с. 13). Электроны, энергия которых лежит в нижней части зоны, ведут себя «нормально»: в ускоряющем поле их кинетическая энергия возрастает, т. е. они движутся от дна в глубь зоны, тем самым эффективная масса таких электронов оказывается положительной, хотя и отличается от значения m . Оба рассмотренных случая можно проследить на зонной диаграмме, учитывающей наличие внешнего поля (см. рис. 1-17, б). При этом следует иметь в виду, что электроны в любой зоне движутся *против* силовых линий поля и что *полная* энергия электрона при его свободном (без столкновений) движении остается постоянной, т. е. электрон перемещается по неизменному («горизонтальному») уровню.

Эффективная масса отражает тот факт, что при наличии внешнего электрического поля электроны в твердом теле движутся не только под действием этого поля (как в случае вакуума), но и под действием периодического *внутреннего* поля, обусловленного атомами решетки. В соотношении между силой и ускорением $F = m (dv/dt)$ можно было бы «усреднить» силу с учетом внешнего и внутреннего полей, но удобнее ввести эффективную массу, а действующей силой считать только *внешнее* поле. Зависимость эффективной массы от энергии и изменение знака массы объясняются волновой природой электрона. Поскольку длина волны электрона $\lambda \sim 1/v$, то при малой энергии (скорости), когда электрон находится в нижней половине зоны, получаем неравенство $\lambda \gg d_0$, где d_0 — межатомное расстояние; тогда электрон слабо взаимодействует с внутренним полем и значение m^* близко к значению m . По мере увеличения энергии длина волны уменьшается ($\lambda \rightarrow d_0$), взаимодействие с решеткой усиливается и эффективная масса растет. В случае $\lambda = d_0$ имеет место своеобразный резонанс, когда внешнее поле «не действует» на электрон, т. е. $m^* = \infty$; это происходит вблизи середины зоны. При дальнейшем повышении энергии электронная волна отражается решеткой, т. е. электрон (при прежнем направлении внешнего поля!) начинает двигаться в обратном направлении; тогда $dv/dt < 0$ и, значит, $m^* < 0$.

Поясним понятие дырки с точки зрения зонной теории. Для этого предположим, что к полупроводнику приложено внешнее напряжение. При наличии электрического поля приходят в движение не только свободные электроны зоны проводимости, но и вся совокупность электронов валентной зоны (поскольку в ней имеются незаполненные верхние уровни, т. е. отсутствует часть электронов с отрицательной эффективной массой). Оказывается, что движение огромного количества электронов валентной зоны с самыми различными эффективными массами эквивалентно движению ограниченного (равного числу незаполненных уровней) количества *фиктивных квазичастиц с положительным зарядом и положительной эффективной массой*. Именно эти квазичастицы получили название *дырок*.

Целесообразность понятия дырок состоит в том, что весьма сложный анализ кинетики электронов в почти полной валентной зоне сводится к сравнительно простому анализу кинетики *дырок* в почти пустой (по отношению к дыркам!) валентной зоне. При этом поведение дырок оказывается аналогичным поведению свободных электронов в почти пустой (по отношению к электронам!) зоне проводимости. Разница сводится к различию эффективных масс обоих типов носителей заряда (табл. 1-1).

Основные параметры некоторых полупроводников

Параметр	Полупроводник			
	Кремний	Германий	Сплав GaAs	Сплав InSb
Заряд ядра	14	32	—	—
Атомный вес	28,1	72,6	—	—
Диэлектрическая проницаемость (отн. ед.)	12	16	11	16
Температура плавления, °C . .	1420	940	1280	520
Коэффициент теплопроводности λ , Вт/(см·°C)	1,2	0,55	—	—
Удельная теплоемкость c , Дж/(г·°C)	0,75	0,41	—	—
Эффективная масса электронов m_n (отн. ед.)	0,33	0,22	0,07	0,013
Эффективная масса дырок m_p (отн. ед.)	0,55	0,39	0,50	0,60
При $T=300$ К:				
Ширина запрещенной зоны φ_z , В	1,11	0,67	1,40	0,18
Эффективная плотность состояний N_c , см ⁻³ . .	$2,8 \cdot 10^{19}$	$1,0 \cdot 10^{19}$	—	—
То же N_v , см ⁻³	$1,0 \cdot 10^{19}$	$0,61 \cdot 10^{19}$	—	—
Подвижность электронов μ_n , см ² /(В·с)	1400	3800	11 000	До 65 000
Подвижность дырок μ_p , см ² /(В·с)	500	1800	450	700
Собственное удельное сопротивление ρ_i , Ом·см	$\sim 2 \cdot 10^5$	~ 60	$\sim 4 \cdot 10^8$	—
Собственная концентрация n_i , см ⁻³	$\sim 2 \cdot 10^{10}$	$2,5 \cdot 10^{13}$	$\sim 1,5 \cdot 10^6$	—
Коэффициент диффузии электронов D_n , см ² /с	36	100	290	До 1750
То же дырок D_p , см ² /с	13	45	~ 12	17
Критическая напряженность поля $E_{крп}$, В/см	2500	900	—	—
То же $E_{крр}$, В/см	7500	1400	—	—
Максимальная скорость электронов $v_{максн}$, см/с	$10 \cdot 10^6$	$6,5 \cdot 10^6$	—	—
То же дырок $v_{макср}$, см/с	$8,0 \cdot 10^6$	$6,0 \cdot 10^6$	—	—

Следует подчеркнуть, что понятие эффективной массы удобно и корректно лишь по отношению к кинетике носителей; его нельзя отождествлять с понятием массы в обычном смысле этого слова. Так, различие эффективных масс отнюдь не характеризует различие силы тяжести [1].

Понятие эффективной массы и ее зависимость от энергии (т. е. от положения электрона в энергетической зоне) наглядно иллюстрируются функциями $W(p)$, где W — энергия и $p = \hbar v$ — импульс электрона (обычно вместо импульса используется про-

Основные физические константы, используемые в теории полупроводников

Основные физические константы	Приближенные значения
Элементарный заряд	$q = 1,6 \cdot 10^{-19}$ Кл
Масса свободного электрона	$m = 9 \cdot 10^{-31}$ кг
Постоянная Планка	$h = 6,6 \cdot 10^{-34}$ Дж·с
Постоянная Больцмана	$k = 1,37 \cdot 10^{-23}$ Дж/°С
Электрическая постоянная (диэлектрическая проницаемость вакуума)	$\epsilon_0 = 9 \cdot 10^{-14}$ Ф/см
Магнитная постоянная (магнитная проницаемость вакуума)	$\mu_0 = 1,25 \cdot 10^{-8}$ Г/см
Число Авогадро (количество атомов в 1 грамме вещества)	$N_0 = 6 \cdot 10^{23}$

порциональный ему квазиимпульс или волновое число $k = \frac{2\pi}{h} p$, где h — постоянная Планка (см. табл. 1-2).

Для получения функции $W(p)$ нужно так или иначе смоделировать кристаллическую решетку, отразив наличие внутренних полей и взаимодействия электронов и атомов [2—5, 7]. В результате зависимость $W(p)$ оказывается набором периодических функций, которые для простейшего одномерного случая показаны на рис. 1-10. Каждый периодический компонент соответствует разрешенной зоне, а «расстояние» между минимумами и максимумами смежных компонентов — запрещенной зоне. С ростом энергии ширина разрешенных зон увеличивается, а ширина запрещенных зон уменьшается, что соответствует интерпретации зон на рис. 1-8, а.

Воспользуемся кривыми $W(p)$ для анализа эффективной массы. Как известно, кинетическая энергия электрона может быть записана в виде

$$W = \frac{mv^2}{2} = \frac{p^2}{2m}.$$

Если $m = \text{const}$ (электрон в вакууме), то $W \sim p^2$, как показано на рис. 1-10 пунктирной параболой. В твердом теле функция $W(p)$ описывается кусочной кривой, показанной жирными линиями; эта кривая отличается от параболы и наличием разрывов, и формой непрерывных участков; следовательно, в твердом теле $m \neq \text{const}$. Дифференцируя дважды функцию $W(p)$, можно

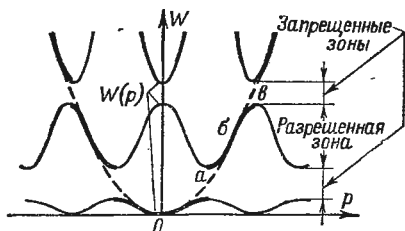


Рис. 1-10. Зависимость энергии электрона от импульса в одномерной кристаллической решетке.

определить эффективную массу электрона как

$$m^* = \frac{1}{d^2W/dp^2} = \left(\frac{h}{2\pi}\right)^2 \frac{1}{d^2W/dk^2}$$

и проследить за ее изменениями с помощью кривых $W(p)$. Для всех зон результат будет один и тот же; ограничимся нижней зоной на рис. 1-10.

Между точками a и b имеем $d^2W/dp^2 > 0$, т. е. $m^* > 0$, что соответствует «обычному» электрону. Между точками b и c имеем $d^2W/dp^2 < 0$, т. е. $m^* < 0$, что соответствует электрону с отрицательной массой. Наконец, в точке перегиба b имеем $d^2W/dp^2 = 0$, т. е. $m^* = \pm \infty$, и, значит (при конечном импульсе p), скорость v обращается в нуль; это соответствует ранее отмеченному «резонансу» между электронной волной и решеткой.

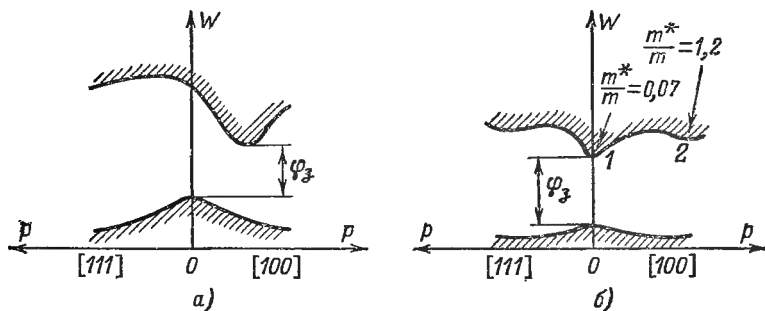


Рис. 1-11. Зонный рельеф кремния (а) и арсенида галлия (б) для двух кристаллографических направлений (ср. с верхней и средней кривыми на рис. 1-10 в пределах 1-го положительного «полупериода»).

В начальной и конечной точках кривой (вблизи минимума и максимума) функция $W(p)$ близка к квадратичной параболе, поэтому на этих участках значения кривизны d^2W/dp^2 почти постоянны, а значит, почти постоянны и эффективные массы, однако они различаются не только по знаку, но и по величине.

В реальном трехмерном случае наглядная интерпретация функции $W(p)$, где p — трехмерный вектор, невозможна, но основные представления, вытекающие из одномерной модели, остаются в силе. Учитывая анизотропию кристалла, следует ожидать (и это подтверждается теорией), что в разных направлениях конфигурации зон, вообще говоря, будут различны. В частности, могут быть различными «расстояния» между разрешенными зонами. Поэтому под шириной запрещенной зоны кристалла понимают «расстояние» между абсолютным минимумом зоны проводимости и абсолютным максимумом смежной (валентной) зоны. Как правило, эти два экстремума соответствуют разным импульсам p , т. е. расположены не «друг под другом» в пространстве импульсов. В этом случае переход электрона из одной зоны в другую должен сопровождаться не только изменением энергии, но и приращением импульса.

Учитывая периодичность функции $W(p)$, ее достаточно анализировать в пределах одного «периода», который в пространстве импульсов носит название зоны Бриллюэна [2, 7]. В частности, в пределах одной зоны Бриллюэна можно

рассмотреть полные рельефы дна зоны проводимости и потолка валентной зоны вдоль того или иного направления (рис. 1-11). Такие рельефы позволяют оценить приращение импульса, необходимое для перехода из зоны в зону (см. выше), выбрать оптимальную плоскость шлифовки кристалла и т. п. Кроме того, из рис. 1-11, б следует, что достаточно энергичные электроны могут перейти из основного минимума 1 в «побочный» минимум 2, где они, как показывает теория, имеют меньшую подвижность. Это обстоятельство приводит к появлению отрицательного сопротивления (например, на кристалле арсенида галлия оказывается возможным строить генераторы СВЧ диапазона — генераторы Ганна [8]).

1-4. ЗОННАЯ СТРУКТУРА ПОЛУПРОВОДНИКОВ

При нулевой абсолютной температуре проводимость в собственном полупроводнике отсутствует, потому что зона проводимости пуста, а валентная зона заполнена. При любой температуре, отличной от нуля, в кристалле появляются фононы, энергетиче-

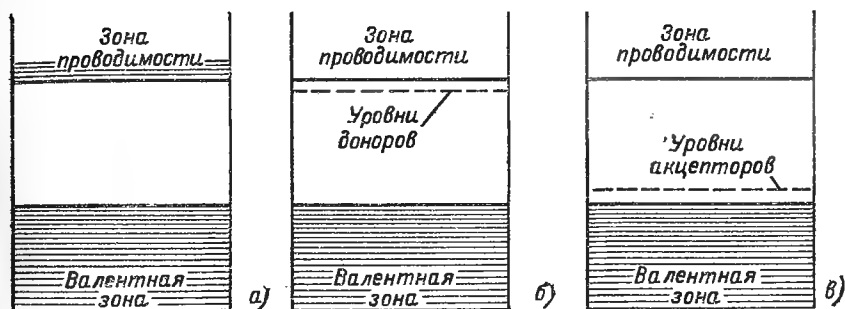


Рис. 1-12. Зонные структуры полупроводников.

а — собственный полупроводник при $T \neq 0$ К; б — электронный полупроводник при $T = 0$ К; в — дырочный полупроводник при $T = 0$ К.

ский спектр которых непрерывен. Фононы с энергией, превышающей ширину запрещенной зоны, переводят некоторые электроны из верхней части валентной зоны в зону проводимости (рис. 1-12, а). В результате в зоне проводимости появляются свободные электроны, а в валентной зоне — незаполненные уровни. Те и другие образуются одновременно и в равных количествах. Теперь электроны обеих зон могут двигаться в электрическом поле, обеспечивая проводимость кристалла.

Перевод электрона из валентной зоны в зону проводимости аналогичен возбуждению электрона в изолированном атоме, т. е. переводу его на один из верхних уровней. Такое возбужденное состояние, как известно, является временным: через некоторое время электрон в атоме возвращается на исходный уровень, а в твердом теле — в валентную зону. В корпускулярном аспекте это возвращение есть рекомбинация электрона и дырки. Среднее время возбужденного состояния электрона (т. е. пребывания его в зоне проводимости) в § 1-2 было названо временем жизни.

С увеличением температуры увеличивается число «энергичных» фононов и возрастает количество электронно-дырочных пар. Соответственно растет собственная проводимость полупроводника.

Зонные диаграммы примесных полупроводников показаны на рис. 1-12, б и в. Так как процентное содержание примесных атомов обычно очень мало, то расстояния между ними относительно велики и их можно рассматривать как изолированные атомы, уровни которых, как известно, не превращаются в зоны. Эти примесные уровни изображены на диаграмме штрихами; каждый штрих соответствует атому примеси.

Донорная примесь характерна тем, что ее валентные уровни располагаются в верхней половине запрещенной зоны (рис. 1-12, б); все эти уровни заполнены при нулевой температуре. Акцепторная примесь характерна тем, что ее валентные уровни располагаются в нижней половине запрещенной зоны (рис. 1-12, в); все эти уровни пусты при нулевой температуре. Переходы электрона с донорного уровня в зону проводимости или из валентной зоны на акцепторный уровень требуют сравнительно небольшой энергии¹. Поэтому при нагреве число фононов, способных ионизировать донор или акцептор и тем самым создать свободный электрон или свободную дырку, будет гораздо больше числа фононов, способных перевести электрон через всю запрещенную зону и образовать электронно-дырочную пару. Соответственно примесная проводимость оказывается гораздо больше собственной. Однако этот вывод относится только к сравнительно низким температурам. Дело в том, что по мере повышения температуры собственная проводимость растет непрерывно, а примесная имеет предел, соответствующий ионизации всех наличных примесных атомов. Таким образом, при достаточно высокой температуре проводимость полупроводника всегда бывает почти собственной.

Если увеличивать концентрацию примесей, то расстояние между примесными атомами уменьшается и их энергетические уровни постепенно превращаются в зоны. Образовавшаяся примесная зона, расширяясь, в конце концов сливается с ближайшей разрешенной зоной кристалла, и получается зонная структура, близкая к металлу. Такой примесный полупроводник называют *вырожденным* или *полуметаллом*.

Обычно в полупроводниках одновременно присутствуют и донорная, и акцепторная примеси, но в разных концентрациях (соответственно N_d и N_a). Рассматривая этот случай, следует учесть, что при нулевой температуре система (кристалл с примесями) должна обладать минимальной энергией. Поэтому электроны с высоких донорных уровней переходят на более низкие — акцеп-

¹ Например, при ширине запрещенной зоны в кремнии около 1 эВ уровни типичных доноров и акцепторов расположены на расстоянии примерно 0,05 эВ от соответствующих разрешенных зон. В германии эти расстояния составляют около 0,01 эВ при ширине запрещенной зоны около 0,7 эВ.

торные, которые, как отмечалось, при нулевой температуре не заполнены. Пусть $N_d > N_a$; тогда при нулевой температуре все акцепторные уровни заполняются электронами (т. е. ионизируются), а число заполненных (неионизированных) донорных уровней уменьшится до $N_d - N_a$ (рис. 1-13, а). При повышении температуры именно эти уровни будут создавать электроны, так как уход электронов с акцепторных уровней требует гораздо большей энергии, соответствующей почти полной ширине запрещенной зоны. Поэтому разность $N_d - N_a$ следует считать *эффективной концентрацией* донорных атомов. Аналогично, если $N_a > N_d$, то при нулевой температуре все доноры ионизируются, а эффективная концентрация акцепторов будет равна $N_a - N_d$ (рис. 1-13, б). В дальнейшем, чтобы не усложнять выражения, мы будем, как правило, понимать под N_d и N_a эффективные концентрации.

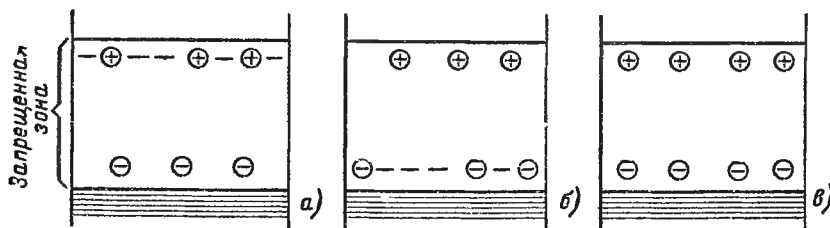


Рис. 1-13. Результирующая концентрация донорных и акцепторных атомов при наличии примесей обоих типов для $T = 0$ К.

а — превалирует донорная примесь; б — превалирует акцепторная примесь; в — количество обеих примесей одинаково.

Если количество обоих типов примеси одинаково ($N_d = N_a$), то эффективные концентрации доноров и акцепторов равны нулю (рис. 1-13, в). Такой полупроводник называется *компенсированным*. Он имеет такую же удельную проводимость, как и собственный, но отличается от последнего (по времени жизни носителей и некоторым другим параметрам), поскольку наличие значительного количества примесей связано с соответствующими искажениями кристаллической решетки.

Некоторые примеси имеют энергетические уровни, расположенные в запрещенной зоне достаточно глубоко, т. е. характеризуются сравнительно высокой энергией ионизации. Эти примеси также играют значительную роль в полупроводниковой технике. Например, если примесные уровни расположены вблизи середины запрещенной зоны (типичный случай — уровни золота в кремнии), то существенно облегчаются как генерация свободных электронов, так и их рекомбинация путем двухступенчатых переходов: из зоны на уровень и с уровня в другую зону. Такие уровни называют *генерационно-рекомбинационными центрами*. Их наличие сильно влияет на время жизни носителей (см. § 1-10). Другим примером могут служить уровни, расположенные вблизи середины верхней

или нижней половины запрещенной зоны. Такие уровни в отличие от генерационно-рекомбинационных центров захватывают носители из ближайшей зоны и через некоторое время отдают их в ту же зону, поскольку «расстояние» до другой зоны значительно больше. Следовательно, уровни этого типа на некоторое время «выводят из игры» часть носителей (электронов или дырок), что сказывается на стабильности, шумах и других свойствах полупроводниковых приборов. Такие уровни называют *ловушками*¹.

В заключение отметим, что многие примеси дают в запрещенной зоне по 2—3 уровня. Например, золото в кремнии дает два уровня: один в середине, другой в нижней половине запрещенной зоны (0,35 В от потолка валентной зоны); последний, несмотря на свое расположение, является донорным. Этот факт, как и другие аналогичные, объясняется возможностью некоторых атомов захватывать не один, а два и больше электронов; тогда, например, однозарядный отрицательный ион, который формально следовало бы отнести к акцепторам, по существу лишен того второго электрона, который он еще мог бы захватить, т. е. должен рассматриваться как донор.

1-5. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ НОСИТЕЛЕЙ В ЗОНАХ ПОЛУПРОВОДНИКА

В теории твердого тела показывается, что энергетические уровни распределены по высоте разрешенной зоны неравномерно: плотность их меняется от границы в глубь зоны. Таким образом, каждому уровню с энергией W соответствует определенная плотность $P(W)$, т. е. число уровней, отнесенное к единице энергии и единице объема твердого тела. Вблизи «дна» и «потолка» каждой из разрешенных зон плотность уровней ($1/\text{Дж} \cdot \text{см}^3$) для узких интервалов энергии dW выражается следующей формулой [1-3]:

$$P(W) = \frac{2}{V\pi} \left(\frac{2\pi m^*}{h^2} \right)^{3/2} \sqrt{W - W_{\text{гр}}}. \quad (1-1a)$$

Здесь h — постоянная Планка (табл. 1-2); m^* — эффективная масса; энергия W отсчитывается от граничного уровня $W_{\text{гр}}$ внутрь зоны. В средних областях зон функция $P(W)$ имеет более сложный вид, но электроны этих областей малосущественны для проводимости.

Величина $W_{\text{гр}}$ является потенциальной энергией электрона или дырки, так как на границах зон скорость частиц, а значит, и их кинетическая энергия равны нулю (см. с. 13); W — полная энергия частиц.

¹ Термин «ловушка» часто относят и к генерационно-рекомбинационным центрам.

Вероятность нахождения электрона на том или ином уровне дается *распределением Ферми — Дирака*

$$F_n(W) = \frac{1}{e^{\frac{W - W_F}{kT}} + 1}, \quad (1-16)$$

где k — постоянная Больцмана (табл. 1-2); T — абсолютная температура; W_F — энергия, называемая *уровнем Ферми*¹.

В дальнейшем будет удобнее выражать энергию не в джоулях, а в электронвольтах или просто в вольтах (численно эти величины одинаковы). Чтобы перейти от одной размерности к другой, достаточно разделить энергии W и kT на элементарный заряд q (см. табл. 1-2). Сделав такую замену в формулах (1-1), получим:

$$P(\varphi) = \frac{2}{\sqrt{\pi}} \left(\frac{2\pi q m^*}{h^2} \right)^{3/2} \sqrt{\varphi - \varphi_{\text{гр}}}; \quad (1-2a)$$

$$F_n(\varphi) = \frac{1}{e^{\frac{\varphi - \varphi_F}{\varphi_T}} + 1}, \quad (1-26)$$

где $P(\varphi)$ — плотность уровней, $1/(\text{В} \cdot \text{см}^3)$; $\varphi = W/q$ — потенциал, характеризующий энергию; φ_F — уровень Ферми в вольтах (потенциал Ферми); φ_T — *температурный потенциал*:

$$\varphi_T = \frac{kT}{q} \approx \frac{T}{11\,600}. \quad (1-3)$$

Название «температурный потенциал» для величины φ_T вполне оправдано, поскольку она имеет размерность напряжения и пропорциональна температуре². Полезно запомнить, что при температуре $T \approx 300 \text{ К}$ (которую мы условно будем называть «комнатной») температурный потенциал

$$\varphi_T(300 \text{ К}) \approx 0,025 \text{ В, или } 25 \text{ мВ.}$$

На рис. 1-14 функции $P(\varphi)$ и $F_n(\varphi)$ показаны на зонной диаграмме собственного полупроводника. Однако ближайшие выводы будут в равной степени относиться к примесным полупроводникам; специфика тех и других будет рассмотрена позже.

В *невырожденных* полупроводниках (см. § 1-6) уровень Ферми φ_F всегда лежит в запрещенной зоне. Глубину его залегания можно характеризовать «расстоянием» от одной из разрешенных зон, выраженным в единицах температурного потенциала φ_T . В боль-

¹ С формальной точки зрения энергия W_F соответствует такому энергетическому уровню, вероятность заполнения которого равна $1/2$. Подробнее уровень Ферми рассматривается в § 1-6.

² С физической точки зрения температурный потенциал есть выраженная в электрических единицах статистическая температура или близкая к ней средняя кинетическая энергия свободного электрона в электронном газе [9].

шинстве случаев уровень φ_F залегает глубоко, т. е. соблюдаясь неравенства¹

$$\varphi_c - \varphi_F \gg \varphi_T; \quad (1-4a)$$

$$\varphi_F - \varphi_v \gg \varphi_T, \quad (1-4б)$$

где φ_c и φ_v — потенциалы «дна» зоны проводимости и «потолка» валентной зоны (индексы происходят от английских названий этих зон: conduction band и valence band).

При температуре $T = 0$ К функция F_n (рис. 1-14) имеет ступенчатый характер, это соответствует уже известным фактам:

валентная зона полностью заполнена ($F_n = 1$), зона проводимости пуста ($F_n = 0$).

При температуре $T \neq 0$ К ступенька функции F_n сглаживается и получается конечная (хотя и крайне малая) вероятность нахождения электронов в зоне проводимости. Одновременно вероятность нахождения электронов в валентной зоне делается немного меньше единицы. В последнем случае удобнее пользоваться вероятностью отсутствия

Рис. 1-14. Плотность уровней энергии, функция вероятности и концентрация носителей в собственном полупроводнике.

электронов на уровнях или, что то же самое, вероятностью наличия дырок:

$$F_p = 1 - F_n = \frac{1}{\frac{\varphi_F - \varphi}{\varphi_T} + 1}. \quad (1-5)$$

Учитывая неравенство (1-4a), нетрудно убедиться, что в зоне проводимости, где $\varphi - \varphi_F > 0$, экспонента в выражении (1-2б) намного превышает единицу и функция F_n упрощается:

$$F_n \approx e^{-\frac{\varphi - \varphi_F}{\varphi_T}}. \quad (1-6a)$$

Аналогично, учитывая неравенство (1-4б), убеждаемся, что и в валентной зоне, где $\varphi - \varphi_F < 0$, экспонента в выражении (1-5) много больше единицы и функция F_p также упрощается:

$$F_p \approx e^{-\frac{\varphi_F - \varphi}{\varphi_T}}. \quad (1-6б)$$

¹ Если обозначить левые части неравенств (1-4) через Δ , то практически эти неравенства будут означать $\Delta \geq 2,3 \varphi_T$, так как тогда $e^{\Delta/\varphi_T} > 10 \gg 1$ и $e^{-\Delta/\varphi_T} < 0,1 \ll 1$, что вполне достаточно для соответствующих пренебрежений и упрощений.

Функции (1-6), которые являются частным случаем распределения Ферми — Дирака (для области энергий, достаточно отличных от φ_F), называются *распределением Максвелла — Больцмана*. Это распределение представляет собой основу теории полупроводников.

Концентрация свободных электронов в зоне проводимости определяется интегралом

$$n = 2 \int_{\varphi_c}^{\infty} P(\varphi - \varphi_c) F_n(\varphi) d\varphi,$$

где подынтегральное выражение есть количество заполненных уровней в элементарном интервале энергий $d\varphi$, расположенном в зоне проводимости, а множитель 2 означает, что на каждом уровне могут (по принципу Паули) находиться два электрона. Подставив (1-2а) и (1-6а) под знак интеграла, после несложных преобразований¹ получим:

$$n = N_c e^{-\frac{\varphi_c - \varphi_F}{\varphi_T}}, \quad (1-7a)$$

где

$$N_c = 2 \left(\frac{2\pi m_n q \varphi_T}{h^2} \right)^{3/2} \approx 0,5 \cdot 10^{16} \left(\frac{m_n}{m} \right)^{3/2} T^{3/2}$$

— *эффективная плотность состояний (на 1 см³) в зоне проводимости*².

Из (1-7а) следует, что N_c есть максимально возможная концентрация электронов в невырожденном полупроводнике, получающаяся при условии $\varphi_F \rightarrow \varphi_c$. По физическому смыслу величина N_c близка к плотности энергетических уровней в зоне проводимости (на 1 см³) в полосе энергий от φ_c до $\varphi_c + \varphi_T$, что можно подтвердить интегрированием функции (1-2а) в соответствующих пределах.

Концентрация свободных дырок в валентной зоне определяется интегралом

$$p = 2 \int_{\varphi_v}^{-\infty} P[-(\varphi - \varphi_v)] F_p(\varphi) (-d\varphi).$$

¹ В процессе преобразований интеграл приводится к виду

$$\int_0^{\infty} e^{-\eta} \sqrt{\eta} d\eta = \sqrt{\pi}/2,$$

где

$$\eta = (\varphi - \varphi_c)/\varphi_T.$$

² Если положить $\varphi_F > \varphi_c$, то из (1-7а) следует качественно правильное соотношение $n > N_c$, характерное для вырожденных полупроводников (см. § 1-6) и металлов. Однако в этом случае для количественной оценки n выражение (1-7а) непригодно, так как оно основано на распределении Максвелла—Больцмана, которое действительно лишь при $\varphi_F < \varphi_c$ [см. (1-4а)].

Подставив сюда (1-2а) и (1-6б), после преобразований получим:

$$p = N_v e^{-\frac{\varphi_F - \varphi_v}{\varphi_T}}, \quad (1-7б)$$

где

$$N_v = 2 \left(\frac{2\pi m_p q \varphi_T}{h^2} \right)^{3/2} \approx 0,5 \cdot 10^{16} \left(\frac{m_p}{m} \right)^{3/2} T^{3/2}$$

— эффективная плотность состояний в валентной зоне (на 1 см³).

Из (1-7б) следует, что N_v есть максимально возможная концентрация дырок в невырожденном полупроводнике, получающаяся при условии $\varphi_F \rightarrow \varphi_v$. По физическому смыслу N_v близка к плотности энергетических уровней в валентной зоне (на 1 см³) в полосе энергий от φ_v до $\varphi_v - \varphi_T$.

Поскольку энергетические зоны, как правило, имеют несколько экстремумов (см. рис. 1-11), то эффективные плотности состояний, как показывает теория, больше приведенных значений: последние следует умножить на число экстремумов. Так, для кремния и германия, у которых имеется соответственно 6 и 4 минимумов в зоне проводимости и один максимум в валентной зоне, получаем данные, приведенные в табл. 1-1.

Заметим, что интегрирование до $\varphi = +\infty$ в случае электронов и до $\varphi = -\infty$ в случае дырок приближенно, поскольку функция (1-2а) действительна только вблизи границ обеих зон. Однако результат получается правильным, так как функции F_n и F_p очень быстро убывают с удалением от зонных границ. Например, на «расстояниях» (4-5) φ_T от границы, когда распределение (1-2а) еще соблюдается, функции F составляют всего около 1% их граничных значений. Заметим еще, что эффективные плотности N_c и N_v различаются, но не сильно (для кремния и германия $N_c/N_v = 2,8$ и 1,7 соответственно). Поэтому иногда полагают $N_v = N_c$.

Наконец, заметим, что экспоненциальные множители в формулах (1-7) по существу являются вероятностями нахождения электронов или дырок в соответствующей зоне, так как величины n/N_c и p/N_v суть отношения фактических концентраций к максимально возможным.

Легко убедиться, что произведение концентраций np не зависит от положения уровня Ферми и определяется только температурой и шириной запрещенной зоны:

$$np = N_c N_v e^{-\varphi_z/\varphi_T} \approx 0,25 \cdot 10^{32} \left(\frac{m_n m_p}{m^2} \right)^{3/2} T^3 e^{-\varphi_z/\varphi_T}, \quad (1-8)$$

где $\varphi_z = \varphi_c - \varphi_v$ — ширина запрещенной зоны.

Значение φ_z несколько уменьшается при увеличении температуры с температурной чувствительностью $3 \cdot 10^{-4}$ и $4 \cdot 10^{-4}$ В/°С соответственно для кремния и германия [10].

Из формул (1-7) легко получить отношение концентраций
следующем виде:

$$\frac{n}{p} = \frac{N_c}{N_v} e^{-\frac{2(\varphi_E - \varphi_F)}{\varphi_T}}, \quad (1-9)$$

где

$$\varphi_E = \frac{\varphi_c + \varphi_v}{2}$$

— потенциал середины запрещенной зоны, который иногда называют *электростатическим потенциалом* полупроводника [2].

1-6. УРОВЕНЬ ФЕРМИ

Уровень Ферми, как мы убедились, играет важную роль в теории полупроводников, а значит, и полупроводниковых приборов. Поэтому имеет смысл уточнить это понятие. Уровень Ферми выше считался известным, с его помощью вычислялись концентрации свободных носителей. На самом же деле уровень Ферми является функцией этих концентраций, а концентрации предварительно оцениваются из тех или иных соображений (см. § 1-7).

Величина φ_F определяется теми интегралами, из которых получены выражения (1-7). Используя (1-2б) и (1-5), эти интегралы легко привести к виду

$$\frac{2}{\sqrt{\pi}} \int_0^{\infty} \frac{V \bar{\eta} d\eta}{e^{\eta - (\chi/\varphi_T)} + 1} = v, \quad (1-10)$$

где для электронов

$$\eta_n = (\varphi - \varphi_c)/\varphi_T; \quad \chi_n = \varphi_F - \varphi_c; \quad v_n = n/N_c,$$

а для дырок

$$\eta_p = (\varphi_v - \varphi)/\varphi_T; \quad \chi_p = \varphi_v - \varphi_F; \quad v_p = p/N_v.$$

Величину χ , определяемую интегральным уравнением (1-10), в статистической физике называют *химическим потенциалом* [5, 9]. Физический смысл этой величины состоит в следующем. Химический потенциал является однозначной функцией концентрации соответствующих частиц. Поэтому наличие разности химических потенциалов означает наличие разности концентраций, а разность концентраций, естественно, вызывает перемещение — *диффузию* частиц в направлении от большей концентрации к меньшей. Таким образом, химический потенциал характеризует возможность диффузии свободных частиц (заряженных или незаряженных), подобно тому как электрический потенциал характеризует возможность *дрейфа* свободных частиц (если они являются носителями заряда) ¹.

¹ В физике полупроводников дрейфом называют движение заряженных частиц в электрическом поле.

Учитывая приведенные выражения для χ_n и χ_p , приходим к выводу, что потенциал Ферми, отсчитанный от границы той или иной зоны (т. е. без учета потенциальной энергии), есть химический потенциал соответствующих носителей. В общем же виде потенциал Ферми есть алгебраическая сумма электрического и химического потенциалов:

$$\varphi_F = \varphi_c + \chi_n; \quad (1-11a)$$

$$\varphi_F = \varphi_v - \chi_p. \quad (1-11b)$$

Отсюда следует еще одно название величины φ_F — *электрохимический потенциал*. Градиент потенциала Ферми, будучи суммой градиентов электрического и химического потенциалов, позволяет одновременно характеризовать оба типа движения носителей — диффузию и дрейф. Как следствие в условиях равновесия, когда направленного движения носителей нет, должно иметь место условие $\text{grad } \varphi_F = 0$, т. е.

$$\varphi_F = \text{const.}$$

Постоянство («горизонтальность») уровня Ферми в равновесной системе является одним из фундаментальных соотношений теории твердого тела. Заметим, что условие $\varphi_F = \text{const}$ не означает постоянства каждого из слагаемых φ_c и χ_n или φ_v и χ_p . Иначе говоря, в равновесной системе могут иметь место градиенты электрического и химического потенциалов и соответственно дрейфовые и диффузионные потоки носителей, но эти потоки должны взаимно уравниваться (*больцмановское равновесие*).

Чтобы выразить химические потенциалы χ_n и χ_p , а вместе с ними и потенциал Ферми через концентрации n и p , нужно решить интегральное уравнение (1-10). В общем виде интеграл в левой части не берется, однако в двух важных частных случаях получаются простые решения [5, 7].

1. Положим $\chi < 0$ и $|\chi| \gg \varphi_T$. Тогда, пренебрегая единицей в знаменателе подынтегрального выражения, легко получаем простое уравнение

$$e^{\chi/\varphi_T} = v,$$

решениями которого будут химические потенциалы

$$\left. \begin{aligned} \chi_n &= \varphi_T \ln \frac{n}{N_c}; \\ \chi_p &= \varphi_T \ln \frac{p}{N_v}. \end{aligned} \right\} \quad (1-12)$$

Учитывая исходные предпосылки относительно χ , приходим к выводу, что полученные решения справедливы при условии $v \ll 1$ (практически при $v \leq 1$). Полупроводники, у которых выполняется это условие, т. е. концентрация свободных носителей меньше эффективной плотности состояний в разрешенной

зоне, называют *невыврожденными*. Для них действительно распределение Максвелла — Больцмана (1-6), которое было использовано при интегрировании. Потенциал Ферми для невырожденных полупроводников имеет вид:

$$\varphi_F = \varphi_c + \varphi_T \ln \frac{n}{N_c}; \quad (1-13a)$$

$$\varphi_F = \varphi_v - \varphi_T \ln \frac{p}{N_v}. \quad (1-13б)$$

Вычитая или складывая выражения (1-13), легко получить соответственно выражения (1-8) и (1-9). Нетрудно заметить также, что у невырожденных полупроводников потенциал Ферми всегда лежит в запрещенной зоне, поскольку логарифмы в обоих выражениях (1-13) отрицательны.

2. Положим $\chi > 0$ и $\chi \gg \varphi_T$. Тогда при $\eta > \chi/\varphi_T$ подынтегральное выражение быстро приближается к нулю, интегрирование в этом диапазоне не имеет смысла. Поэтому примем в качестве верхнего предела интегрирования $\eta = \chi/\varphi_T$. С учетом исходных предпосылок относительно χ можно считать знаменатель подынтегрального выражения равным единице в диапазоне $0 \leq \eta < \chi/\varphi_T$.

Выполнив интегрирование, получим простое уравнение:

$$\frac{4}{3\sqrt{\pi}} \left(\frac{\chi}{\varphi_T} \right)^{3/2} = v,$$

решениями которого будут химические потенциалы:

$$\chi_n = \left(\frac{3\sqrt{\pi}}{4} \right)^{2/3} \varphi_T \left(\frac{n}{N_c} \right)^{2/3}; \quad \chi_p = \left(\frac{3\sqrt{\pi}}{4} \right)^{2/3} \varphi_T \left(\frac{p}{N_v} \right)^{2/3}.$$

Эти решения действительны при условии $v \gg 1$ (практически при $v > 3$), так как с самого начала принято $\chi \gg \varphi_T$. Полупроводники, у которых соблюдается условие $v > 1$, т. е. концентрация свободных носителей существенно превышает эффективную плотность состояний в разрешенной зоне, называют *вырожденными* или *полуметаллами* (см. § 1-4). Для них распределение Максвелла — Больцмана недействительно и в случае сильного вырождения заменяется ступенчатой функцией ($F \approx 1$ при $\eta < \chi/\varphi_T$ и $F \approx 0$ при $\eta > \chi/\varphi_T$), которая была использована при интегрировании. Критерии вырождения ($v > 1$) имеют вид:

$$n > N_c = 2 \left(\frac{2\pi m_n q \varphi_T}{h^2} \right)^{3/2}; \quad (1-14a)$$

$$p > N_v = 2 \left(\frac{2\pi m_p q \varphi_T}{h^2} \right)^{3/2}. \quad (1-14б)$$

Легко убедиться, что потенциал Ферми для вырожденных полупроводников лежит в н у т р и соответствующей разрешенной зоны, поскольку химические потенциалы χ_n и χ_p положительны. В частности, это относится к металлам (см. рис. 1-9).

Термин «уровень Ферми» обычно используется для равновесного состояния системы, в котором значения (1-13а) и (1-13б) совпадают. В неравновесном состоянии значения (1-13а) и (1-13б), вообще говоря, различны, тогда их называют соответственно *квазиуровнями Ферми* для электронов и для дырок: φ_{F_n} и φ_{F_p} [2].

1-7. КОНЦЕНТРАЦИЯ НОСИТЕЛЕЙ

В собственном полупроводнике концентрации свободных электронов и дырок одинаковы: $n = p$. Тогда из формул (1-9) и (1-13) следует, что при любой температуре уровень Ферми собственного полупроводника расположен вблизи середины запрещенной зоны (см. рис. 1-14), т. е. $\varphi_{F_i} \approx \varphi_E^*$. Подставляя $n = p$ в формулу (1-8), легко получаем концентрации свободных электронов и дырок в собственном полупроводнике (индекс i от английского слова intrinsic — настоящий):

$$n_i = p_i = \sqrt{N_c N_v} e^{-\frac{\varphi_g}{2\varphi_T}} = 0,5 \cdot 10^{16} \left(\frac{m_n m_p}{m^2} \right)^{3/4} T^{3/2} e^{-\frac{\varphi_g}{2\varphi_T}}. \quad (1-15)$$

Равенству концентраций n_i и p_i соответствует на рис. 1-14 идентичность кривых, характеризующих концентрацию свободных носителей в соответствующей разрешенной зоне.

Зависимость собственных концентраций n_i и p_i от температуры очень сильна и обусловлена в основном изменением температурного потенциала в показателе экспоненты, а не степенным множителем $T^{3/2}$. Столь же сильно зависит собственная концентрация от ширины запрещенной зоны при данной температуре. Так, сравнительно небольшое различие в значении φ_g у германия и кремния (0,67 и 1,11 В) приводит к различию собственных концентраций при комнатной температуре более чем на 3 порядка (см. табл. 1-1). Сравнивая (1-8) и (1-15), соотношение (1-8) можно записать в более компактной форме:

$$np = n_i^2, \quad (1-16)$$

которую и будем использовать в дальнейшем.

Соотношение (1-16) говорит о том, что увеличение концентрации одного типа носителей сопровождается уменьшением концентрации другого типа носителей. Так, для электронных полупроводников, у которых $n \gg n_i$, имеем $p \ll n_i$, а для дырочных полупроводников, у которых $p \gg n_i$, имеем $n \ll n_i$.

* Строго говоря,

$$\varphi_{F_i} = \varphi_E + \frac{\varphi_T}{2} \ln \frac{N_v}{N_c} \approx \varphi_E + 0,4\varphi_T,$$

однако эта поправка практически несущественна.

Используя формулы (1-16) и (1-7) и полагая $N_c = N_v$, нетрудно выразить концентрации n и p через собственную концентрацию n_i :

$$n = n_i e^{-\frac{\varphi_E - \varphi_F}{\varphi_T}}; \quad (1-17a)$$

$$p = n_i e^{-\frac{\varphi_F - \varphi_E}{\varphi_T}}. \quad (1-17б)$$

Отсюда легко получить потенциал Ферми в двух формах:

$$\varphi_F = \varphi_E + \varphi_T \ln \frac{n}{n_i}; \quad (1-18a)$$

$$\varphi_F = \varphi_E - \varphi_T \ln \frac{p}{n_i}. \quad (1-18б)$$

С помощью выражений (1-18) легко еще раз убедиться в том, что у электронных полупроводников, у которых $n \gg n_i$ и $p \ll n_i$,

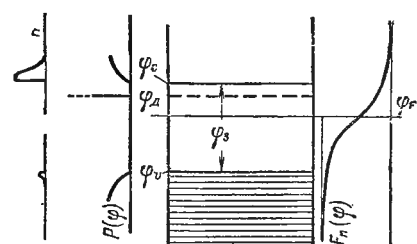


Рис. 1-15. Плотность уровней энергии, функция вероятности и концентрация носителей в примесном полупроводнике n -типа.

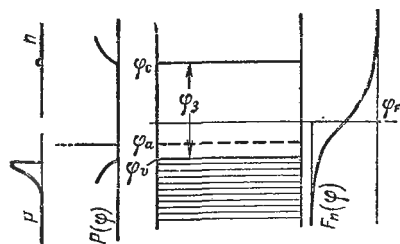


Рис. 1-16. Плотность уровней энергии, функция вероятности и концентрация носителей в примесном полупроводнике p -типа.

уровень Ферми лежит в верхней половине запрещенной зоны (рис. 1-15), а у дырочных полупроводников, у которых $p \gg n_i$ и $n \ll n_i$, — в нижней половине запрещенной зоны (рис. 1-16) *.

Для того чтобы определить потенциал φ_F по формулам (1-13) или (1-18), нужно знать концентрации свободных носителей. При оценке значений n и p используют условие нейтральности (точнее, квазинейтральности) полупроводника. Это важное условие обосновано в работах [2, 11] (см. также § 1-12) и формулируется следующим образом: *в однородном полупроводнике не может быть существенных uncompensated объемных зарядов ни в равновесном состоянии, ни при наличии тока*. Поэтому в общем виде

* Поскольку энергетические уровни примесных атомов совпадают, плотность этих уровней на диаграммах характеризуется бесконечно узкими и бесконечно высокими «пиками», площадь которых равна концентрации доноров или акцепторов. Математически такие «пики» описываются δ -функцией.

условие нейтральности для единичного объема записывают так

$$p + N_d^* - (n + N_a^*) = 0, \quad (1-19a)$$

где N_d^* , N_a^* — концентрации ионизированных доноров и акцепторов¹. Уравнение (1-19a) говорит о том, что концентрация частиц, несущих положительный заряд (дырки и ионизированные доноры), равна концентрации частиц, несущих отрицательный заряд (электроны и ионизированные акцепторы).

Для электронных полупроводников, не содержащих акцепторов,

$$n = N_d^* + p. \quad (1-19b)$$

Для дырочных полупроводников, не содержащих доноров,

$$p = N_a^* + n. \quad (1-19v)$$

В случае равновесного состояния к уравнению (1-19b) легко прийти путем следующих рассуждений. Разделим концентрацию свободных электронов на две части:

$$n = n_d + n_v,$$

где концентрация n_d обусловлена электронами доноров, а концентрация n_v — электронами валентной зоны. Так как уход каждого электрона из валентной зоны сопровождается образованием дырки, то $n_v = p$. Так как уход электрона с донорного уровня сопровождается образованием ионизированного донора, то $n_d = N_d^*$, где N_d^* — эффективная концентрация ионизированных доноров. С учетом этих соотношений равенство $n = n_d + n_v$ переходит в условие нейтральности (1-19b). Аналогичные рассуждения для дырочного полупроводника приводят к условию (1-19v).

Прежде чем использовать уравнения (1-19), необходимо определить концентрации ионизированных примесей.

Значение N_d^* можно найти из следующих соображений. Вероятность заполнения донорного уровня φ_d равна $F_n(\varphi_d)$. Тогда $1 - F_n(\varphi_d)$ есть вероятность отсутствия электрона на донорном уровне, т. е. вероятность ионизации донора. Умножая эту вероятность на N_d , находим искомую концентрацию ионизированных доноров²:

$$N_d^* = \frac{N_d}{\frac{\varphi_F - \varphi_d}{2e^{\varphi_T}} + 1}. \quad (1-20a)$$

Подставляя в (1-20a) значение φ_F из (1-13a), после небольших преобразований получаем:

$$N_d^* = \frac{N_d}{2 \frac{n}{N_c} e^{\varphi_{cl}/\varphi_T} + 1}. \quad (1-20b)$$

¹ Ионизация примеси может быть либо следствием перехода носителя в одну из разрешенных зон, либо захвата его примесью другого типа (см. рис. 1-13).

² В качестве функции F_n использовано общее выражение (1-26), поскольку расположение уровня Ферми относительно уровня доноров не ограничено условиями типа (1-4) и, следовательно, использование упрощенного выражения (1-6a) недопустимо. Коэффициент 2 при экспоненциальном члене является результатом более строгого анализа (см. [5]).

Здесь $\varphi_{cd} = \varphi_c - \varphi_d$ — «разность» между донорными уровнями и зоной проводимости, которое можно назвать *потенциалом ионизации доноров*.

Учитывая, что вероятность ионизации акцептора есть вероятность заполнения акцепторного уровня, т. е. $F_n(\varphi_a)$, нетрудно получить аналогичные выражения для концентрации ионизированных акцепторов N_a^* . Для этого в формулах (1-20) достаточно заменить N_d на N_a , $\varphi_F - \varphi_d$ на $\varphi_a - \varphi_F$, n на p , N_c на N_v и φ_{cd} на φ_{va} , где $\varphi_{va} = \varphi_a - \varphi_v$ — потенциал ионизации акцепторов.

Перейдем непосредственно к оценке концентраций свободных носителей. Рассмотрим сначала электронные полупроводники.

В типичном электронном полупроводнике выполняется неравенство $n \gg p$. Кроме того, в рабочем диапазоне температур донорные атомы практически полностью ионизированы, т. е. $N_d^* \approx N_d$. Тогда из соотношения (1-19б) получаем концентрацию свободных электронов:

$$n = N_d, \quad (1-21a)$$

которая, как видим, определяется концентрацией примеси и, следовательно, не зависит от температуры. С помощью (1-16) легко получаем концентрацию свободных дырок:

$$p = n_i^2 / N_d, \quad (1-21б)$$

которая согласно (1-15) очень сильно — экспоненциально — зависит от температуры. Наконец, из (1-18a) или (1-13a) находим уровень Ферми в типичном электронном полупроводнике:

$$\varphi_F = \varphi_E + \varphi_T \ln \frac{N_d}{n_i} = \varphi_c + \varphi_T \ln \frac{N_d}{N_c}. \quad (1-21в)$$

Выражение (1-21в) позволяет сделать вывод, что уровень Ферми лежит тем выше, чем больше концентрация доноров и чем ниже температура.

Простые формулы (1-21) широко используются на практике. Однако следует иметь в виду, что они действительны в ограниченном температурном диапазоне: с понижением температуры степень ионизации доноров уменьшается и принятое равенство $N_d^* = N_d$ становится менее строгим. С повышением температуры увеличивается концентрация собственных носителей и постепенно нарушается принятое неравенство $n \gg p$ (электронный полупроводник превращается в собственный).

Обозначим через T_1 и T_2 нижнюю и верхнюю границы температурного диапазона, в котором действительны формулы (1-21); назовем T_1 *температурой ионизации примеси*, а T_2 — *критической температурой* полупроводника.

Положим, что полная ионизация соответствует условию $N_d^* = 0,9 N_d$. Тогда из выражения (1-20б), в котором можно принять $n = N_d^* = 0,9 N_d$, находим температуру ионизации в следующей

невяной форме ¹:

$$\varphi_T(T_1) \ln \frac{N_c(T_1)}{16N_d} = \varphi_{сн}. \quad (1-22a)$$

Это трансцендентное уравнение решается методом последовательных приближений. Подставляя для германия $\varphi_{сн} = 0,01$ В и полагая $N_d = 10^{16}$, получаем $T_1 \approx 64$ К. Для кремния при $\varphi_{сн} = 0,05$ В и той же концентрации N_d получается $T_1 \approx 145$ К. С увеличением концентрации доноров температура ионизации возрастает, но практически она всегда меньше нижней границы рабочего диапазона полупроводниковых приборов (-60°C или 213 К).

Положим, что превращение примесного полупроводника в собственный соответствует условию $n = 100 p$. Тогда из (1-16) легко находим: $n = 10 n_i$; $p = n_i/10$. Если подставить эти значения в (1-19б), принять $N_d^* = N_d$, а концентрацию n_i выразить с помощью (1-15), считая $N_v = N_c$, то после преобразований получим критическую температуру в форме, аналогичной (1-22a) ²:

$$\varphi_T(T_2) \ln \frac{10N_c(T_2)}{N_d} = \frac{\varphi_3}{2}. \quad (1-22б)$$

Для германия с концентрацией доноров $N_d = 10^{16}$ получается $T_2 \approx 360$ К, т. е. около 90°C . Для кремния при той же концентрации примеси критическая температура существенно выше и составляет около 250°C . *Это является одним из важнейших преимуществ кремния как материала для полупроводниковых приборов.* С увеличением концентрации доноров критическая температура растет.

Если выразить концентрацию примесей через удельное сопротивление (см. § 1-8), то после преобразований формулу (1-22б) можно записать в полуэмпирической форме:

$$T_2 (^\circ\text{C}) = 273 \left(\frac{M}{m + \lg \rho} - 1 \right), \quad (1-23)$$

где для кремния $M = 10$; $m = 4,53$; для германия $M = 6,6$; $m = 4,63$, а ρ выражено в Ом·см [12].

Оценим концентрацию носителей и уровень Ферми за границами температурного диапазона $T_1 - T_2$.

Если в выражении (1-19б) положить $N_d^* = N_d$; $p = n_i^2/n$ и решить получающееся квадратное уравнение относительно n , то концентрация электронов в области *высоких температур* ($T > T_2$) примет вид:

$$n = \sqrt{\left(\frac{N_d}{2}\right)^2 + n_i^2} + \frac{N_d}{2}. \quad (1-24a)$$

¹ Допущение $n = N_d^*$ вытекает из (1-19б), поскольку в области низких температур неравенство $n \gg p$ сохраняется и даже усиливается.

² Допущение $N_d^* = N_d$ оправдано тем, что в области высоких температур степень ионизации доноров согласно (1-20а) возрастает. В формулу (1-22б) подставляют значения φ_3 с учетом температурной чувствительности (см. с. 26).

Концентрация дырок в этой области получится из соотношения (1-19б) при $N_d^* = N_d$:

$$p = \sqrt{\left(\frac{N_d}{2}\right)^2 + n_i^2} - \frac{N_d}{2}. \quad (1-24б)$$

Подставляя найденные значения концентраций в формулы (1-18), легко получить уровень Ферми в области высоких температур:

$$\varphi_F = \varphi_E + \varphi_T \ln \left[\sqrt{\left(\frac{N_d}{2n_i}\right)^2 + 1} + \frac{N_d}{2n_i} \right]. \quad (1-24в)$$

Очевидно, что выражения (1-24) являются обобщениями выражений (1-21).

Если в выражении (1-19б) положить $p = 0$, подставить в него N_d^* из (1-20б) и решить получающееся квадратное уравнение относительно n , то концентрация электронов в области низких температур ($T \lesssim T_1$) имеет вид:

$$n = \left[\sqrt{\left(\frac{N_c}{4}\right)^2 + \frac{N_c N_d}{2} e^{\varphi_{cd}/\varphi_T}} - \frac{N_c}{4} \right] e^{-\varphi_{cd}/\varphi_T}. \quad (1-25а)$$

В области сверхнизких температур, когда можно пренебречь членами $N_c/4$ [этому случаю соответствует пренебрежение единицей в знаменателе (1-20б), т. е. условие $N_d^* \ll N_d$], формула (1-25а) упрощается и переходит в следующую:

$$n = \sqrt{\frac{N_c N_d}{2}} e^{-\frac{\varphi_{cd}}{2\varphi_T}}. \quad (1-25б)$$

Такому значению концентрации n соответствует согласно (1-13а) уровень Ферми:

$$\varphi_F = \varphi_c - \frac{\varphi_{cd}}{2} + \frac{\varphi_T}{2} \ln \frac{N_d}{2N_c}. \quad (1-25в)$$

Нетрудно убедиться, что последнее слагаемое в формуле (1-25в) стремится к нулю, когда $T \rightarrow 0$, т. е. уровень Ферми при нулевой температуре расположен посередине между уровнем доноров и дном зоны проводимости. Полезно также заметить, что формула (1-25б) имеет большое структурное сходство с формулой (1-15): по существу плотность уровней N_v заменена концентрацией N_d , а ширина запрещенной зоны $\varphi_z = \varphi_c - \varphi_v$ — шириной «запрещенной зоны» $\varphi_{cd} = \varphi_c - \varphi_d$. Такой результат вполне естествен, поскольку в области сверхнизких температур собственные носители практически отсутствуют и «поставщиком» электронов являются только донорные уровни, играющие роль валентной зоны.

Мы рассмотрели концентрацию носителей для отдельных участков температурного диапазона. Общее выражение для концентрации n и соответственно для величин p и φ_F можно найти, подставив в (1-19б) значение N_d^* из (1-20б) и значение p из (1-16). Получающееся при этом полное кубическое уравнение относительно n в принципе решается, но решение слишком громоздко и имеет чисто теоретическое значение.

Случай дырочного полупроводника, в котором $p \gg n$, нет необходимости рассматривать столь же подробно. Главные особенности акцепторной примеси видны из рис. 1-16. Если акцепторы полностью ионизированы и температура ниже критической, то вместо формул (1-21) получаем аналогичные соотношения, характер-

ные для ярко выраженного дырочного полупроводника:

$$p = N_a; \quad (1-26a)$$

$$n = n_i^2 / N_a; \quad (1-26б)$$

$$\varphi_F = \varphi_E - \varphi_T \ln \frac{N_a}{n_i} = \varphi_v - \varphi_T \ln \frac{N_a}{N_v}. \quad (1-26в)$$

Как видим, уровень Ферми лежит тем ниже, чем больше концентрация акцепторов и чем меньше температура.

Остаются в силе понятия температуры ионизации и критической температуры. Для их определения в формулах (1-22) следует заменить N_d на N_a , N_c на N_v и $\varphi_{сж}$ на $\varphi_{ва}$.

Аналогичная замена параметров в формулах (1-24) и (1-25) позволяет найти концентрации в области высоких и низких температур и затем определить уровень Ферми.

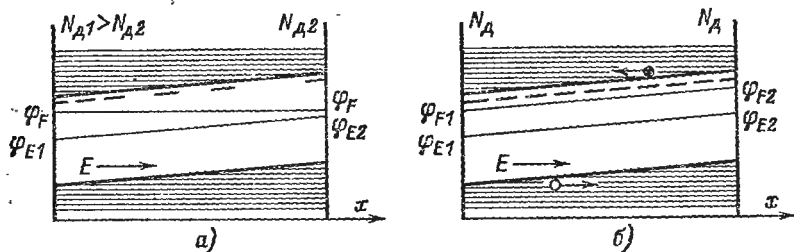


Рис. 1-17. Зонные диаграммы неоднородного полупроводника (а) и однородного полупроводника при наличии внешнего электрического поля (б).

Все предыдущие зонные диаграммы соответствовали *однородным полупроводникам*, в которых примеси распределены совершенно равномерно. Разумеется, однородный полупроводник является некоторой идеализацией. Более того, часто специально создают неоднородность внутри кристалла в виде градиента концентрации примесей, что придает полупроводнику свойства, необходимые для ряда приборов. Посмотрим, каковы особенности *неоднородных полупроводников*.

Пусть, например, в полупроводнике типа *n* концентрация доноров изменяется от N_{d1} до $N_{d2} < N_{d1}$. Согласно (1-21в) разность $\varphi_F - \varphi_E$ уменьшается с уменьшением концентрации N_d . Поскольку в равновесной системе уровень Ферми во всех ее частях одинаков, зонная диаграмма должна иметь такой вид, как на рис. 1-17, а. Электростатический потенциал φ_E вдоль неоднородного полупроводника меняется. Следовательно, в *неоднородных полупроводниках имеются внутренние электрические поля*, в которых возможен дрейф носителей. Однако в отсутствие внешнего поля дрейфовые потоки носителей равны противоположно направленным диффузионным потокам тех же носителей, обусловленным градиен-

том их концентрации (в общем случае — градиентом химического потенциала). Поэтому результирующий поток отсутствует и соблюдается больцмановское равновесие (см. § 1-6).

Для сравнения на рис. 1-17, б показана зонная диаграмма однородного полупроводника при наличии внешнего электрического поля (напряженность $E = d\varphi_E/dx$ та же, что и на рис. 1-17, а). Если в силу условия квазинейтральности принять концентрации носителей неизменными вдоль оси x , то будут неизменными и химические потенциалы, т. е. «расстояния» уровня φ_F от краев разрешенных зон. Тогда согласно (1-18) получаем $d\varphi_F/dx = = d\varphi_E/dx$, т. е. имеет место градиент уровня Ферми, как и должно быть при нарушении равновесия (внешнее электрическое поле вызывает протекание тока)¹. Наличие градиента уровня Ферми обуславливает принципиальное отличие данного случая от предыдущего (рис. 1-17, а) несмотря на внешнее сходство «перекошенных» зонных диаграмм.

Заметим, что наличие внутреннего электрического поля, вообще говоря, означает нарушение условия квазинейтральности, однако если поле почти постоянное, то объемные заряды не существенны.

Все сказанное действительно и для полупроводника типа p с учетом специфики расположения уровня Ферми. Легко, например, убедиться, что при убывании концентрации акцепторов слева направо зоны искривляются не вверх, как на рис. 1-17, а, а вниз.

Зонные диаграммы неоднородных полупроводников строятся по формулам (1-21в) и (1-26в).

1-8. ПОДВИЖНОСТЬ НОСИТЕЛЕЙ

Удельная проводимость любого тела зависит не только от концентрации носителей, но и от их подвижности в электрическом поле. Подвижность носителей, по определению, есть их средняя направленная скорость в электрическом поле с напряженностью 1 В/см. Соответственно дрейфовую скорость можно записать в виде $v_{др} = \mu E$.

Постоянство дрейфовой скорости носителей в однородном поле ($E = \text{const}$) специфично для твердого тела, где имеются различные «препятствия» движению. В вакууме, где таких препятствий нет, движение заряженных частиц в однородном поле равноускоренное, т. е. понятие дрейфовой скорости отсутствует. Направленное движение носителей в твердом теле под действием поля сочетается с их хаотическим (тепловым) движением. Последнее характери-

¹ Отрицательно заряженные электроны, которые двигаются против поля, можно представить себе на диаграмме в виде тяжелых шариков, «скатывающихся» по наклонному дну зону проводимости. Положительно заряженные дырки, которые двигаются по полю, можно представить себе как легкие пузырьки, «всплывающие» по наклонному потолку «жидкой» валентной зоны.

зудается средней тепловой скоростью

$$v_T = \sqrt{3kT/m^*}, \quad (1-27)$$

зависящей от температуры. Если принять $T = 300$ К и $m^* = m$, то $v_T \approx 10^7$ см/с.

При не слишком сильных электрических полях дрейфовая скорость намного меньше тепловой или, как говорят, температура носителей определяется температурой кристаллической решетки. В этих условиях подвижность можно выразить формулой

$$\mu = \frac{q}{m^*} t_{cp} = \frac{q}{m^*} \frac{l_{cp}}{v_T}, \quad (1-28)$$

где t_{cp} — среднее время свободного пробега частицы; l_{cp} — средняя длина ее свободного пробега (для кремния и германия $t_{cp} = (2 \div 5) 10^{-13}$ с и $l_{cp} = 0,02 \div 0,05$ мкм).

Величины t_{cp} и l_{cp} характеризуют частоту столкновений носителей с теми или иными «препятствиями». В результате таких столкновений происходит изменение скорости и направления движения носителей, т. е. их *рассеяние*. Теория показывает [7, 10], что при нулевой абсолютной температуре в идеальной кристаллической решетке рассеяние не имеет места. Иначе говоря, атомы решетки как таковые не являются препятствием на пути движения носителей. Истинными препятствиями являются лишь колеблющиеся атомы решетки (в корпускулярной интерпретации — фононы), а также атомы примесей и дефекты структуры.

В области очень низких температур интенсивность тепловых колебаний мала, количество фононов незначительно и рассеяние обусловлено главным образом дефектами решетки и нейтральными примесями (нейтральность примесей является следствием низкой температуры, меньшей температуры ионизации).

В области обычных рабочих температур рассеяние обусловлено главным образом фононами и ионизированными примесями. В этой области подвижность μ можно представить состоящей из двух компонентов μ_L и μ_I , связанных с результирующей подвижностью формулой¹

$$\frac{1}{\mu} = \frac{1}{\mu_L} + \frac{1}{\mu_I}. \quad (1-29)$$

Очевидно, что результирующая подвижность определяется наименьшим из компонентов μ_L и μ_I .

Анализ показывает [2, 4, 7, 10], что при решеточном рассеянии $l_{cp} \sim 1/(m^*)^2 T$. Тогда из (1-28) с учетом (1-27) следует соотношение

$$\mu_L \sim (m^*)^{-5/2} T^{-3/2}. \quad (1-30a)$$

¹ Индексы L и I происходят от начальных букв в словах *Lattice* — решетка и *Ion* — ион.

При и о н н о м рассеянии получается соотношение ¹:

$$\mu_I \sim (m^*)^{-1/2} T^{3/2} N^{-1}, \quad (1-306)$$

где N — концентрация ионизированной однозарядной примеси.

Из соотношений (1-30) видно, что подвижности μ_L и μ_I , а следовательно, и результирующая подвижность μ обратно пропорциональны эффективной массе. Этим, в частности, объясняется различие подвижностей электронов и дырок. Как правило, $\mu_n > \mu_p$. Если воспользоваться общепринятым обозначением

$$b = \mu_n / \mu_p, \quad (1-31)$$

то получим: для кремния $b = 2,8$ и для германия $b = 2,1$.

Соотношения (1-30) позволяют исследовать зависимость подвижности от температуры и концентрации примеси.

В обычном температурном диапазоне полупроводниковых приборов и при не очень высокой концентрации примеси выполняется условие $\mu_L < \mu_I$. В этом случае согласно (1-29) $\mu \approx \mu_L$, т. е. результирующая подвижность определяется решеточным рассеянием и зависимость подвижности от температуры согласно (1-30а) должна иметь вид:

$$\mu = \mu_0 \left(\frac{T_0}{T} \right)^c, \quad (1-32)$$

где μ_0 — подвижность при температуре T_0 , например комнатной (см. табл. 1-1); $c = 3/2$. Однако реальные значения показателя c у большинства полупроводников отличаются от теоретического. Для электронного и дырочного кремния $c_n = 2,6$; $c_p = 2,3$; для германия $c_n = 1,66$; $c_p = 2,33$; для арсенида галлия $c_n = 1,0$; $c_p = 2,1$.

Из (1-32) следует, что подвижность при $T \rightarrow 0$ должна неограниченно расти. На самом деле этого не происходит, так как при достаточно низких температурах величина μ_I становится меньше μ_L , т. е. согласно (1-29) превалирует ионное рассеяние, при котором $\mu \sim T^{3/2}$ [см. (1-306)]. Типичные кривые $\mu(T)$ показаны на рис. 1-18 для рабочего диапазона температур.

Зависимость подвижности от концентрации примеси обусловлена ионным рассеянием. Действительно, согласно (1-306), рост концентрации N приводит к уменьшению компонента μ_I , а значит, в той или иной степени — и к уменьшению результирующей подвижности. Типичные кривые $\mu(N)$ показаны на рис. 1-19. Как видим, при $N < 10^{15} \text{ см}^{-3}$ подвижность можно считать постоянной;

¹ Соотношение (1-306) — упрощенный вариант формулы Конуэлл—Вайскопфа [4, 7], в которой опущен знаменатель

$$\ln \left[1 + \left(\frac{12\pi e_0 e k T}{q^2 N^{1/3}} \right)^2 \right],$$

несущественный при обычных температурах и концентрациях.

в интервале $10^{15} < N < 10^{16} \text{ см}^{-3}$ она уменьшается в 1,5—2 раза, а при еще больших концентрациях падает в несколько раз. Учитывая, что сильная (гиперболическая) зависимость $\mu_i(N)$ «скрадывается» постоянной составляющей μ_L , следует ожидать (и это подтверждается на практике), что результирующая подвижность μ зависит от N слабее, чем по закону N^{-1} . Соответственно можно

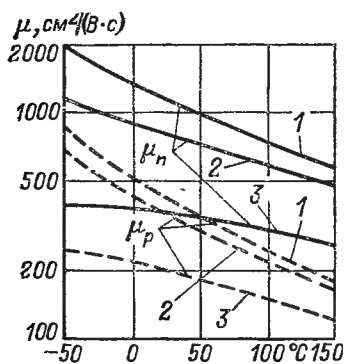


Рис. 1-18. Зависимость подвижности носителей в кремнии от температуры при разной концентрации примеси.

1 — $N = 10^{16}$; 2 — $N = 10^{17}$; 3 — $N = 10^{18}$.

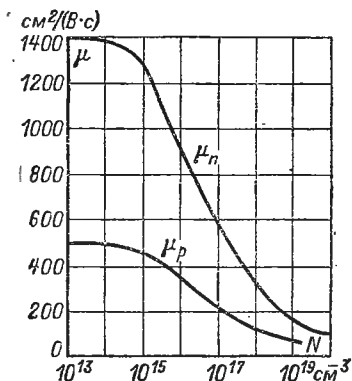


Рис. 1-19. Зависимость подвижности носителей в кремнии от концентрации примеси при $T = 300 \text{ К}$.

предложить полуэмпирическую зависимость $\mu \sim N^{-1/3}$ [13]; тогда для достаточно больших концентраций ($N > N_0 = 10^{15} \div 10^{16} \text{ см}^{-3}$)

$$\mu = \mu_0 \left(\frac{N_0}{N} \right)^{1/3}, \quad (1-33a)$$

где подвижность μ_0 соответствует концентрации N_0 .

Иногда удобнее оказывается чисто эмпирическая аппроксимация [14, с. 147]:

$$\mu = \mu_0 - \Delta\mu \lg \frac{N}{N_0}, \quad (1-336)$$

в которой коэффициент $\Delta\mu$ есть изменение подвижности на декаду приращения концентрации.

Наличие зависимости $\mu(N)$, в частности, означает, что в неоднородном полупроводнике, у которого концентрация примеси различна в разных его частях, подвижность является функцией координаты, т. е. меняется вдоль пути движущегося носителя. Это обстоятельство иногда приходится учитывать при анализе и расчете полупроводниковых приборов.

Важную роль играет зависимость подвижности от напряженности электрического поля. Если напряженность превышает кри-

тическое значение $E_{кр}$ (см. табл. 1-1), то подвижность

$$\mu = \mu_0 \left(\frac{E_{кр}}{E} \right)^{1/2}, \quad (1-34)$$

где $E > E_{кр}$, а μ_0 — подвижность при $E = E_{кр}$. Режим $E > E_{кр}$ носит название *сверхкритического*.

Критическая напряженность поля не является строго определенной величиной; она лишь приблизительно соответствует условию, когда направленная (дрейфовая) скорость носителей делается сравнимой с их хаотической (тепловой) скоростью, например $\mu E_{кр} = 0,1v_T$. Однако даже такое приближенное условие позволяет констатировать, что критическая напряженность, учитывая (1-27), (1-32) и (1-33), является функцией температуры ($E_{кр} \sim T^{c+1/2}$), концентрации ($E_{кр} \sim N^{1/3}$) и начальной подвижности ($E_{кр} \sim 1/\mu_0$). Последняя зависимость, в частности, объясняет меньшие значения $E_{кр}$ у германия по сравнению с кремнием.

Сравнимые значения дрейфовой и тепловой скоростей позволяют говорить о том, что «суммарная» скорость носителей в сверхкритическом режиме повышается, а значит, их температура оказывается соответственно выше окружающей. Носители с повышенной температурой, т. е. с энергией, сравнимой или превышающей тепловую энергию $\frac{3}{2} kT$, называют *горячими*¹.

Горячие носители, взаимодействуя (сталкиваясь) с имеющимися акустическими фононами, сами «подогревают» решетку, порождая новые, более энергичные оптические фононы (см. сноску на с. 8). Порождение оптических фононов [10] приводит к новому явлению — *насыщению скорости*. А именно, при достаточно сильных полях [обычно при $E \geq (4 \div 5) E_{кр}$] скорость носителей перестает зависеть от напряженности поля; т. е. получается условие $v = \mu E = \text{const}$.

Такой результат равносителен соотношению $\mu \sim E^{-1}$ вместо соотношения $\mu \sim E^{-1/2}$, лежащего в основе (1-34). Таким образом, полупроводникам свойственно понятие *максимальной скорости* носителей $v_{\text{макс}}$. Эта скорость соответствует энергии $W_{\text{о.ф.}}$, необходимой для возбуждения оптического фонона. Как только носитель в интервале между столкновениями накопит энергию $W_{\text{о.ф.}}$ и скорость $v_{\text{макс}}$, он тотчас же теряет их на образование оптического фонона и опять начинает разгоняться с «нуля». Поэтому значение $v_{\text{макс}}$ можно оценить из элементарного соотношения $\frac{1}{2} m^* v^2 = W_{\text{о.ф.}}$, где энергия $W_{\text{о.ф.}}$, как показывает теория, близка к тепловой энергии «холодной» решетки $\frac{3}{2} kT$. Получающееся значение $v_{\text{макс}}$ при-

¹ Поскольку энергия kT характеризуется температурным потенциалом (1-3), который при $T = 300$ К составляет около 25 мВ, можно считать, что при комнатной температуре горячими являются носители с энергией 0,03 эВ и выше.

мерно составляет 10^7 см/с. Более точные значения приведены в табл. 1-1.

Зависимость μ (E) оказывает непосредственное влияние на вольт-амперную характеристику полупроводника. Например, в простейшем случае (однородный кристалл с постоянным сечением)

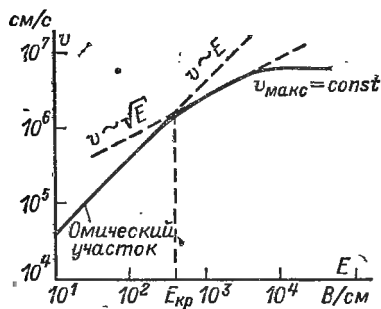


Рис. 1-20. Зависимость скорости носителей от напряженности электрического поля.

ток I пропорционален скорости v , а напряжение U — напряженности электрического поля E , поэтому форма кривой $I(U)$ такая же, как кривой $v(E)$ (рис. 1-20)*. На начальном участке соблюдается закон Ома, поскольку $\mu = \text{const}$ и, следовательно, $v \sim E$. На последующих двух участках этот закон все больше нарушается, вплоть до насыщения тока. Значит, при достаточно больших напряжениях (когда поле $E > E_{\text{кр}}$) полупроводник ведет себя как нелинейный резистор.

1-9. УДЕЛЬНАЯ ПРОВОДИМОСТЬ И УДЕЛЬНОЕ СОПРОТИВЛЕНИЕ

В общем виде удельная проводимость выражается следующим образом¹:

$$\sigma = \frac{1}{\rho} = qn\mu_n + qp\mu_p, \quad (1-35)$$

где μ_n , μ_p — подвижности электронов и дырок; ρ — удельное сопротивление.

* Действительно, при постоянном сечении и постоянном удельном сопротивлении значение U определяется элементарным соотношением $U = Ew$, где w — длина образца. Что касается значения I , то оно (при постоянном сечении) пропорционально плотности тока, определяемой соотношением

$$j = \lambda v,$$

где λ — концентрация подвижных зарядов. В однородном кристалле $\lambda = \text{const}$ и, следовательно, $I \sim v$.

¹ Это и последующие выражения основаны на соотношении для плотности тока, приведенном в предыдущей сноске. В самом деле, учитывая два типа носителей, можно записать:

$$j = \lambda_n v_n + \lambda_p v_p,$$

где $\lambda_n = qn$ и $\lambda_p = qp$. Подставляя $v_n = \mu_n E$ и $v_p = \mu_p E$ и сравнивая полученное выражение с дифференциальной формой закона Ома ($j = \sigma E$), приходим к формуле (1-35) и ее частным случаям.

Для частных случаев собственного, электронного и дырочного полупроводников получаем соответственно:

$$\sigma_i = \frac{1}{\rho_i} = qn_i(\mu_n + \mu_p); \quad (1-36a)$$

$$\sigma_n = \frac{1}{\rho_n} \approx qn\mu_n; \quad (1-36б)$$

$$\sigma_p = \frac{1}{\rho_p} \approx qp\mu_p. \quad (1-36в)$$

Если в формуле (1-35) концентрацию p выразить через n с помощью соотношения (1-16), то, как легко убедиться путем дифференцирования, минимальное значение удельной проводимости

$$\sigma_{\min} = 2qn_i \sqrt{\mu_n \mu_p} \quad (1-37)$$

получается при концентрациях

$$n = n_i / \sqrt{b}; \quad p = n_i \sqrt{b},$$

где b — отношение подвижностей (1-31). Сравнивая (1-37) и (1-36a), нетрудно показать, что $\sigma_{\min} < \sigma_i$, т. е. минимальная удельная проводимость свойственна не собственному, а «слегка» примесному полупроводнику. Если $b > 1$ (как обычно и бывает), то значение $\sigma_{\min}$ соответствует дырочной проводимости. Для германия и кремния значение $\sigma_{\min}$ на 7 и 12% соответственно меньше, чем σ_i .

Температурную зависимость удельной проводимости или удельного сопротивления можно получить, зная температурную зависимость концентрации носителей (§ 1-7) и их подвижности (§ 1-8).

В широком температурном диапазоне зависимость $\sigma(T)$ удобно изображать в полулогарифмическом масштабе, откладывая по оси абсцисс «обратную температуру» $1/T$ (рис. 1-21). В этом случае для собственного полупроводника получается прямая, наклон которой пропорционален $0,5 \varphi_3$. Этот результат следует из выражения (1-36a), если подставить в него (1-15) и (1-32).

Для примесных полупроводников зависимость $\sigma(T)$ получается сложнее. В области очень низких температур (большие значения $1/T$), когда степень ионизации примеси мала, получается прямая с наклоном, пропорциональным $0,5 \varphi_{сд}$ или $0,5 \varphi_{ва}^*$. Так

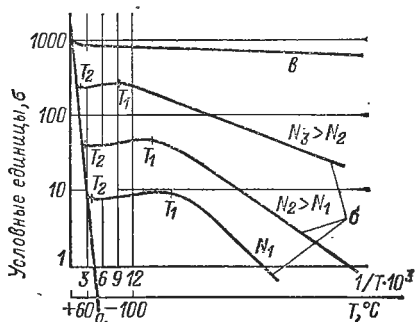


Рис. 1-21. Зависимость удельной проводимости полупроводника от температуры.

а — собственный полупроводник; б — примесные полупроводники; в — полуметалл.

* Действительно, если, например, уровни доноров почти все заполнены, то совокупность этих уровней образует своего рода валентную зону, отделенную от зоны проводимости «расстоянием» $\varphi_{сд}$. Последнее играет роль ширины запрещенной зоны, и вместо наклона $0,5 \varphi_3$ получается наклон $0,5 \varphi_{сд}$.

как $\varphi_{сд}, \varphi_{за} \ll \varphi_{в}$, то этот наклон значительно меньше, чем в собственном полупроводнике. По мере ионизации примесей наклон кривой уменьшается, и при полной ионизации получается почти горизонтальный участок. Начиная с этой температуры (T_1) и до критической температуры (T_2) концентрация основных носителей практически постоянна [см. (1-21a) и (1-26a)]. Следовательно, на этом участке проводимость меняется так же, как подвижность, т. е. по закону (1-32). При дальнейшем повышении температуры ($T > T_2$) проводимость переходит в собственную и резко возрастает.

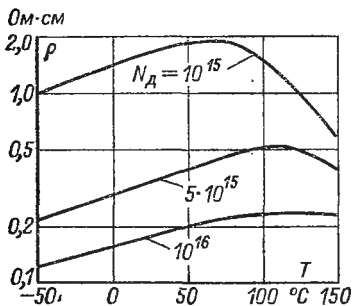


Рис. 1-22. Зависимость удельного сопротивления германия типа n от температуры при различных концентрациях доноров.

При очень большой концентрации примесей полупроводник превращается в полуметалл (см. § 1-4 и 1-6) с очень большой проводимостью, мало зависящей от температуры.

На рис. 1-22 для наглядности показана зависимость удельного сопротивления от температуры в линейном масштабе и в более узком диапазоне, характерном для применения в полупроводниковой технике.

В условиях полной ионизации примесей и докритической температуры, когда $n \approx N_d$ или $p \approx N_a$, формулы (1-36б) и (1-36в) позволяют

оценить концентрацию примесей по значению удельного сопротивления¹. Например, при $\rho = 1$ Ом·см в кремнии типа n получается $N_d = 5 \cdot 10^{15}$ см⁻³. Для кремния типа p при том же значении ρ концентрация акцепторов N_a будет больше примерно втрое из-за меньшего значения μ_p . По той же причине концентрация примесей в германии примерно в 3 раза меньше, чем в кремнии, при одном и том же ρ .

1-10. РЕКОМБИНАЦИЯ НОСИТЕЛЕЙ

Общие сведения. Процессы генерации и рекомбинации носителей заряда неотъемлемы друг от друга и в то же время противоположны по содержанию. Генерация является ведущим началом в этом единстве и связана с воздействием таких внешних факторов, как нагрев, освещение или облучение. Рекомбинация представляет собой внутреннюю реакцию системы на появление и возрастание числа носителей. Именно рекомбинация, противодействуя накоплению носителей, обуславливает их равновесные концентрации, рассмотренные в § 1-7; в частности, она лежит в основе фундаментального

¹ Поскольку при таких оценках считается известной подвижность, которая, однако, сама является функцией концентрации, более точные измерения концентраций N_d и N_a осуществляются с помощью эффекта Холла [7, 10, 15].

постоянства (1-8). Не меньшую роль играет рекомбинация в нестационарных процессах; благодаря ей свободные носители имеют конечное время жизни, а время жизни характеризует скорость изменения концентраций электронов и дырок при быстрых изменениях внешних факторов. Поэтому изучение механизма рекомбинации и ее количественных закономерностей необходимо для понимания и использования многих важнейших явлений в полупроводниках и полупроводниковых приборах.

Непосредственная рекомбинация свободного электрона со свободной дыркой — сравнительно редкое событие [10]: время жизни носителей, вычисленное исходя из непосредственной рекомбинации, на несколько порядков больше наблюдаемых значений¹. Поэтому главную роль следует отвести механизму рекомбинации с помощью *центров рекомбинации* [17], которые часто называют «ловушками» (см. конец § 1-4). Напомним, что ловушке свойственны энергетические уровни, расположенные глубоко в запрещенной зоне, близко к ее середине. Переход электрона из зоны проводимости на уровень ловушки и затем в валентную зону гораздо более вероятен, чем непосредственный переход через запрещенную зону при непосредственной рекомбинации. На рис. 1-23 показаны две возможные последовательности процессов при рекомбинации на ловушках.

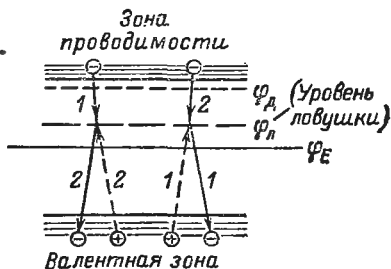


Рис. 1-23. Возможные варианты рекомбинации носителей на ловушках.

1, 2 — последовательность этапов;
— — — эквивалентное перемещение дырок.

Поскольку ловушка в равной степени облегчает переход электронов как из зоны проводимости в валентную зону, так и в обратном направлении, она представляет собой не только центр рекомбинации, но и центр генерации носителей, как и должно быть из общих соображений о равновесии.

С физической точки зрения понятие ловушек столь же широко, как и понятие примесей; это могут быть и посторонние атомы, и различные подвижные частицы, и дефекты кристаллической решетки.

В частности, если полупроводник представляет собой поликристалл (см. сноску¹ на с. 6), то внутренние грани составляющих его

¹ Этот вывод не является универсальным. Однако он справедлив для кремния, у которого экстремумы зон в пространстве импульсов не совпадают (рис. 1-11, а), т. е. рекомбинирующие частицы должны не только «встретиться», но и иметь определенные импульсы, что маловероятно. У таких материалов, как арсенид галлия, у которого экстремумы зон совпадают (рис. 1-11, б), непосредственная рекомбинация гораздо более вероятна и, поскольку она часто сопровождается выделением кванта света, используется в таких приборах, как световые диоды [16].

микрорекристаллов образуют своеобразные дефекты решетки, что приводит к уменьшению времени жизни носителей. Для изготовления полупроводниковых диодов и транзисторов используют монокристаллы с регулярной структурой, у которых время жизни обычно лежит в пределах 10—100 мкс.

До сих пор мы имели в виду рекомбинацию в объеме полупроводника и объемное время жизни. Однако не менее, а часто более существенное значение имеет рекомбинация на поверхности. Это объясняется тем, что дефекты решетки неизбежны в первую очередь на поверхности кристалла, где нарушена симметрия связей атомов, а также наиболее вероятно наличие разного рода пленок (адсорбированные газы, влага, окислы и т. п.). Поэтому поверхность полупроводника представляет собой особую, весьма активную область, содержащую большое число энергетических уровней. Поверхностные уровни, расположенные в запрещенной зоне, могут играть роль ловушек, и тогда поверхность становится областью интенсивной рекомбинации и генерации носителей. Обычно эти поверхностные процессы характеризуют не временем жизни, а *скоростью поверхностной рекомбинации* S , измеряемой в сантиметрах в секунду. Этот коэффициент является сложной функцией геометрии полупроводника, состояния его поверхности и подвижности носителей. С физической стороны коэффициент S соответствует средней скорости носителей, с которой они движутся к поверхности, где происходит непрерывное «уничтожение» их. Величина S лежит в весьма широких пределах: от 100 до 10^4 см/с и более.

В зависимости от конфигурации образца и расположения токоподводящих электродов превалирует тот или иной вид рекомбинации — объемный или поверхностный. Роль поверхностной рекомбинации возрастает с увеличением отношения площади образца к его объему, т. е. при прочих равных условиях с уменьшением размеров образца. Однако на практике разделение рекомбинации на две составляющие далеко не всегда необходимо. Поэтому для анализа и расчета полупроводниковых приборов чаще всего используют единый параметр — так называемое *эффективное время жизни* τ , которое характеризует совместное влияние объемной и поверхностной рекомбинаций и определяется соотношением

$$\frac{1}{\tau} = \frac{1}{\tau_v} + \frac{1}{\tau_s}, \quad (1-38)$$

где τ_v и τ_s — *объемное* и *поверхностное* времена жизни. Величина τ_s является функцией скорости поверхностной рекомбинации S .

Эффективное время жизни в полупроводниковых диодах и транзисторах обычно составляет 0,1—2 мкс, но в ряде случаев может быть на порядок больше и меньше этих значений.

Рассмотрим теперь количественные закономерности процессов рекомбинации и генерации носителей. Сначала для простоты будем предполагать непосредственную рекомбинацию, а позднее учтем практически более важный ловушечный механизм.

Равновесное состояние. В равновесном состоянии полупроводника процессы генерации — рекомбинации подчиняются *закону действующих масс*. Пусть имеется электронный полупроводник, в котором, как известно, наличие свободных электронов обусловлено двумя факторами: ионизацией донорных атомов и ионизацией атомов основного материала (см. с. 32). Запишем соответствующие обратимые реакции:

(свободный электрон) + (ионизированный донор) $\rightleftharpoons$ (нейтральный донор);

(свободный электрон) + (свободная дырка) $\rightleftharpoons$ (связанный электрон валентной зоны).

Обозначим концентрации частиц, участвующих в этих реакциях, через n_0 , N_d^* , $N_d - N_d^*$, p_0 , N_v — величины, известные из предыдущих параграфов ¹.

Тогда согласно закону действующих масс количественные характеристики записанных реакций должны выражаться формулами [9]

$$\frac{n_0 N_d^*}{(N_d - N_d^*)} = K_1(T); \quad (1-39a)$$

$$\frac{n_0 p_0}{N_v} = K_2(T), \quad (1-39б)$$

где $K_1(T)$ и $K_2(T)$ — так называемые *коэффициенты действующих масс*, зависящие от температуры.

Из выражения (1-39a), в частности, следует, что примесные атомы принципиально не могут быть ионизированы полностью, хотя разность $N_d - N_d^*$ может быть очень малой. Выражение (1-39б) равносильно соотношению (1-8), которое было получено с учетом конкретных законов распределения и в котором поэтому коэффициент $K_2(T)$ «расшифрован»:

$$K_2(T) = N_v e^{-\varphi_3/\varphi T}.$$

Произведение $n_0 p_0$ в формуле (1-39б) по смыслу соответствующей реакции характеризует процесс рекомбинации, а произведение $K_2(T) N_v$ — процесс генерации. Таким образом, скорость рекомбинации пропорциональна произведению концентраций рекомбинирующих частиц. Это обстоятельство вполне естественно, так как рекомбинация электрона тем вероятнее, чем больше электронов в данном объеме и чем больше дырок, с которыми он может рекомбинировать.

Исходя из приведенных соображений, можно записать условие равновесия (1-39б) в форме

$$r(n_0 p_0) = g_0, \quad (1-40)$$

¹ Концентрации с индексом 0 здесь и в дальнейшем соответствуют равновесному состоянию. Для концентрации связанных электронов вместо $N_v - p_0$ использовано значение N_v , поскольку $p_0 \ll N_v$.

где левая часть есть скорость рекомбинации (r — коэффициент рекомбинации), а $g_0 = rK_2N_v$ — скорость генерации. Параметры r и g_0 не зависят от концентрации свободных носителей.

Величина rp_0n_0 есть количество актов рекомбинации в единице объема и в единицу времени, а n_0 — количество электронов в единице объема. Следовательно, rp_0 есть вероятность рекомбинации одного электрона в единицу времени¹, а обратная величина будет средним интервалом между актами рекомбинации, т. е. средним временем жизни электронов:

$$\tau_n = \frac{1}{rp_0}. \quad (1-41a)$$

Аналогичные рассуждения приводят к выражению для среднего времени жизни дырок:

$$\tau_p = \frac{1}{rn_0}. \quad (1-41b)$$

Формулы (1-41) можно получить из (1-40) непосредственно, если воспользоваться следующим определением времен жизни:

$$\left. \begin{aligned} \tau_n &= \frac{n_0}{R_0}; \\ \tau_p &= \frac{p_0}{R_0}, \end{aligned} \right\} \quad (1-42)$$

где $R_0 = rn_0p_0$ — скорость рекомбинации носителей в равновесном состоянии.

Как видно из (1-41) и (1-42), равновесные времена жизни электронов и дырок в общем случае резко различны. Это объясняется тем, что скорости рекомбинации обоих типов носителей одинаковы (так как они рекомбинируют парами), а концентрации n_0 и p_0 могут различаться на много порядков.

С использованием времени жизни условие равновесия (1-40) запишется следующим образом:

$$\frac{n_0}{\tau_n} = \frac{p_0}{\tau_p} = g_0. \quad (1-43)$$

Еще одну форму записи того же уравнения (1-40) можно получить, используя вероятности нахождения электронов и дырок в той или иной зоне. Рассмотрим отдельно записи левой и правой частей (1-40) при таком методе.

Пусть вероятность наличия электрона в верхней части валентной зоны есть F_{nv} . Тогда вероятность наличия дырки в той же части зоны будет $1 - \bar{F}_{nv}$.

¹ Отсюда следует, что коэффициент рекомбинации r есть вероятность захвата электрона в единицу времени при единичной концентрации дырок. Такой вывод подтверждается структурой коэффициента r [5]:

$$r = \sigma_{\text{эфф}} v_T,$$

где $\sigma_{\text{эфф}}$ — эффективное сечение захвата электрона одним из атомов, содержащих дырки; v_T — средняя тепловая скорость электронов (1-27).

Эта вероятность ¹ равна отношению фактической концентрации дырок p_0 к концентрации N_v :

$$1 - \bar{F}_{nv} = p_0 / N_v.$$

Если из этого равенства выразить концентрацию p_0 и подставить ее в левую часть (1-40), то величину $n_0 p_0$ можно записать как скорость рекомбинации электронов:

$$R_{n0} = C_{nR} n_0 (1 - \bar{F}_{nv}). \quad (1-44a)$$

Коэффициент $C_{nR} = r N_v$ по аналогии с величиной $r p_0$, рассмотренной при выводе (1-41a), есть вероятность захвата электрона в единицу времени при $p_0 = N_v$, т. е. при условии, что верхняя часть валентной зоны протяженностью Φ_T пуста. Обратная величина $C_{nR}^{-1} = \tau_{n0}$ есть среднее время жизни электронов при том же условии, т. е. в ярко выраженном (идеальном) дырочном полупроводнике. Время жизни τ_{n0} в отличие от τ_n не зависит от концентрации подвижных носителей, т. е. является параметром материала.

Используя вероятность наличия электрона в нижней части зоны проводимости $\bar{F}_{nc} = n_0 / N_c$, можно записать левую часть (1-40) как скорость рекомбинации дырок:

$$R_{p0} = C_{pR} p_0 \bar{F}_{nc}, \quad (1-44b)$$

где $C_{pR} = r N_c$ — вероятность захвата дырки при $n_0 = N_c$, т. е. при условии полностью заполненной нижней части зоны проводимости протяженностью Φ_T . Обратная величина $C_{pR}^{-1} = \tau_{p0}$ есть среднее время жизни дырок в ярко выраженном (идеальном) электронном полупроводнике, не зависящее от концентрации подвижных носителей.

Поскольку величины N_v и N_c одного порядка (см. § 1-5), времена жизни τ_{n0} и τ_{p0} тоже близки друг к другу, хотя и могут различаться в несколько раз.

Правую часть (1-40) можно представить в такой же форме, как (1-44), если рассматривать генерацию носителей как «обратную рекомбинацию». Соответственно скорость генерации можно записать либо как скорость генерации электронов:

$$g_{n0} = C_{ng} N_v (1 - \bar{F}_{nc}) \approx C_{ng} N_v, \quad (1-45a)$$

либо как скорость генерации дырок:

$$g_{p0} = C_{pg} N_c \bar{F}_{nv} \approx C_{pg} N_c. \quad (1-45b)$$

Коэффициенты C_{ng} и C_{pg} — вероятности генерации соответствующих носителей при условии, что зона проводимости пуста, а валентная зона заполнена. Значения C_{ng} и C_{pg} близки друг к другу.

Уравнение (1-40) можно записать в разных формах, приравнивая выражения (1-44) и (1-45). При этом устанавливается однозначная связь между коэффициентами C_g и C_R :

$$\left. \begin{aligned} C_{ng} &= C_{nR} \frac{n'}{N_v}; \\ C_{pg} &= C_{pR} \frac{p'}{N_c}, \end{aligned} \right\} \quad (1-46a)$$

где

$$\left. \begin{aligned} n' &= N_c e^{-\Phi_3 / \Phi_T}; \\ p' &= N_v e^{-\Phi_3 / \Phi_T}. \end{aligned} \right\} \quad (1-46b)$$

¹ Буква F с черточкой означает «интегральную» вероятность, поскольку речь идет о наличии электронов или дырок на любом из уровней зоны (в интервале Φ_T , отсчитанном от граничного уровня), тогда как буква F без черточки означает вероятность наличия носителя на определенном уровне зоны (см. § 1-5).

Выражения (1-46) позволяют записать скорость генерации в виде

$$g_0 = r N_c N_v e^{-\varphi_3 / \varphi T}.$$

Таким образом, сущность «вероятностного метода» сводится к следующему: скорость рекомбинации записывается как произведение концентрации рекомбинирующих носителей на вероятность наличия другого типа носителей в соответствующей зоне; коэффициент пропорциональности C_R равен вероятности рекомбинации (в единицу времени), когда условия рекомбинации оптимальны, т. е. когда верхняя часть валентной зоны пуста (в случае рекомбинации электронов), либо когда нижняя часть зоны проводимости полностью заполнена (в случае рекомбинации дырок). Аналогично записываются скорости генерации носителей. Обе скорости можно приравнять и установить связь между коэффициентами пропорциональности.

Неравновесное состояние. В неравновесном состоянии, когда скорости генерации и рекомбинации неодинаковы, происходит накопление (или рассасывание) неравновесных носителей со скоростью, определяемой разностью $g - R$:

$$\frac{dn}{dt} = g_n - r(np); \quad (1-47a)$$

$$\frac{dp}{dt} = g_p - r(np). \quad (1-47b)$$

Неравновесные концентрации n и p можно представить в виде

$$\left. \begin{aligned} n &= n_0 + \Delta n; \\ p &= p_0 + \Delta p, \end{aligned} \right\} \quad (1-48)$$

где Δn , Δp — избыточные концентрации. Знак приращений Δn и Δp соответствует знаку приращения скорости генерации, если эту скорость представить в виде

$$g = g_0 + \Delta g. \quad (1-49)$$

Сохранение нейтральности полупроводника (см. § 1-7) предполагает равенство избыточных концентраций и их производных:

$$\Delta n = \Delta p; \quad (1-50a)$$

$$\frac{dn}{dt} = \frac{dp}{dt}. \quad (1-50b)$$

Условие (1-50b) сразу приводит к выводу, что в формулах (1-47) должно иметь место равенство $g_n = g_p$, т. е. скорости генерации электронов и дырок одинаковы. Соответственно уравнения (1-47a) и (1-47b) оказываются идентичными и не имеет смысла анализировать отдельно поведение избыточных электронов и дырок, так как функции $\Delta n(t)$ и $\Delta p(t)$ должны совпадать. Поэтому ниже рассматривается только поведение электронов.

Подставим в (1-47a) выражения (1-48) и (1-49), учтем (1-40) и согласно (1-50a) заменим Δp на Δn . Далее положим $\Delta n \ll n_0 + p_0$ (чтобы избежать нелинейности уравнения) и выразим концентрации p_0 и n_0 через времена жизни τ_n и τ_p с помощью (1-41). После этого

уравнение (1-47а) можно привести к виду

$$\frac{d(\Delta n)}{dt} = \Delta g - \frac{\Delta n}{\tau}. \quad (1-51a)$$

Если положить $\Delta g = 0$, получится частный случай, который называется *уравнением рассасывания носителей*:

$$\frac{d(\Delta n)}{dt} = -\frac{\Delta n}{\tau}. \quad (1-51b)$$

В уравнениях (1-51) через τ обозначено неравновесное время жизни и з б ы т о ч н ы х носителей:

$$\tau = \left(\frac{1}{\tau_n} + \frac{1}{\tau_p} \right)^{-1} = \frac{\tau_n \tau_p}{\tau_n + \tau_p}; \quad (1-52)$$

в условиях нейтральности неравновесное время жизни о д и н а к о в о для электронов и дырок. Очевидно, что τ ближе к м е н ь ш е м у из двух равновесных времен жизни τ_n и τ_p . Например, для электронного полупроводника, у которого $n_0 \gg p_0$ и $\tau_n \gg \tau_p$ [см. (1-41)], получаем $\tau \approx \tau_p$; для дырочного полупроводника, наоборот, $\tau \approx \tau_n$. Иначе говоря, изменение и з б ы т о ч н ы х концентраций происходит с постоянной времени, близкой к равновесному времени жизни н е о с н о в н ы х носителей.

Если $\Delta g = \text{const}$ (ступенчатое изменение скорости g), то решением уравнения (1-51a) будет экспоненциальная функция

$$\Delta n(t) = \Delta g \tau + [\Delta n(0) - \Delta g \tau] e^{-\frac{t}{\tau}}, \quad (1-53a)$$

где $\Delta n(0)$ — начальное значение избыточной концентрации.

Решением уравнения рассасывания (1-51b) также является экспоненциальная функция

$$\Delta n(t) = \Delta n(0) e^{-\frac{t}{\tau}}. \quad (1-53b)$$

Уравнение рассасывания и его решение лежат в основе измерений времени жизни [18]. Так, из (1-51b) получается общее о п р е д е л е н и е:

$$\tau = -\frac{\Delta n}{dn/dt}, \quad (1-54)$$

где в правой части стоят мгновенные значения величин. Однако практически удобнее измерять значение Δn в р а з н ы е моменты времени t_1 и t_2 ; тогда из (1-53b) следует:

$$\tau = \frac{t_2 - t_1}{\ln \frac{\Delta n(t_1)}{\Delta n(t_2)}}.$$

Рекомбинация на ловушках. При ловушечном механизме рекомбинации (см. рис. 1-23) уравнения (1-51) и (1-53) остаются в силе,

но время жизни τ имеет другой вид. Анализ ловушечной рекомбинации удобно провести «вероятностным методом», т. е. с помощью выражений типа (1-44) и (1-45).

Будем считать, что концентрация ловушек N_t мала по сравнению с концентрацией примесей ¹, а их энергетические уровни одинаковы и характеризуются потенциалом φ_t . Вероятность заполнения уровня φ_t определяется формулой (1-26) при $\varphi = \varphi_t$, т. е. имеет вид ²:

$$F_t = \frac{1}{e^{\frac{\varphi_t - \varphi_F}{\varphi_T}} + 1}. \quad (1-55)$$

Согласно рис. 1-23 уровни ловушек в процессе рекомбинации играют роль «валентной зоны» по отношению к зоне проводимости и роль «зоны проводимости» по отношению к валентной зоне. Поэтому в формулах (1-44) естественно заменить $\bar{F}_{nv}$ и $\bar{F}_{nc}$ на F_t , а N_v и N_c (в выражениях для C_R) на N_t . В процессе генерации совокупность занятых ловушечных уровней играет роль «валентной зоны» по отношению к зоне проводимости, а совокупность пустых ловушечных уровней — роль «зоны проводимости» по отношению к валентной зоне. Поэтому в формулах (1-45) заменим плотность N_v на $N_t F_t$, а плотность N_c на $N_t (1 - F_t)$.

С учетом сделанных замен скорости накопления (1-47) будут выражаться следующим образом:

$$\frac{dn}{dt} = C_{ng} N_t F_t - C_{nR} n (1 - F_t); \quad (1-56a)$$

$$\frac{dp}{dt} = C_{pg} N_t (1 - F_t) - C_{pR} p F_t, \quad (1-56b)$$

где $C_{nR} = r_n N_t$ и $C_{pR} = r_p N_t$.

Рассматривая стационарный режим ($dn/dt = 0$; $dp/dt = 0$; $n = n_0$; $p = p_0$), используя для n_0 и p_0 выражения (1-7) и подставляя (1-55), можно из (1-56) получить соотношения

$$\left. \begin{aligned} C_{ng} &= C_{nR} \frac{n_t}{N_t}; \\ C_{pg} &= C_{pR} \frac{p_t}{N_t}, \end{aligned} \right\} \quad (1-57a)$$

по форме такие же, как при непосредственной рекомбинации (1-46a). Однако в отличие от концентраций n' и p' концентрации n_t и p_t

¹ В том случае, когда концентрация N_t сравнима с концентрацией примесных атомов или больше ее, становится недействительным условие (1-50a), так как ловушки при этом не только играют роль центров рекомбинации, но и вносят заметный вклад в общее число свободных носителей наряду с основной примесью ([10], с. 288). Индекс t — от слова trap (ловушка).

² Приближение (1-6a) для функции F_t неправомерно, так как уровни φ_t и φ_F часто близки друг к другу и даже совпадают.

выражаются формулами

$$\left. \begin{aligned} n_t &= N_c e^{-\frac{\Phi_c - \Phi_t}{\Phi_T}}; \\ p_t &= N_v e^{-\frac{\Phi_t - \Phi_v}{\Phi_T}} \end{aligned} \right\} \quad (1-576)$$

у которых в показателе экспоненты стоит не «расстояние» между разрешенными зонами, как в (1-466), а «расстояние» между одной из зон и уровнем ловушек, поскольку именно оно согласно принятой методике анализа играет роль «запрещенной зоны» при ловушечном механизме рекомбинации.

Сравнивая формулы (1-576) с формулами (1-7), приходим к выводу, что концентрации n_t и p_t соответствуют тому случаю, когда уровень ловушек совпадает с уровнем Ферми. Произведение этих концентраций удовлетворяет условию (1-16)

$$n_t p_t = n_0 p_0 = n_i^2, \quad (1-58a)$$

а их отношение описывается формулой

$$\frac{n_t}{p_t} = \frac{N_c}{N_v} e^{-2(\Phi_F - \Phi_t)/\Phi_T}, \quad (1-586)$$

аналогичной (1-9), но с заменой уровня Φ_F на Φ_t .

Важной особенностью ловушечной рекомбинации является *неравенство скоростей генерации электронов и дырок*, а следовательно, и скоростей их рекомбинации¹. С физической точки зрения причина этих неравенств состоит в наличии двух возможных последовательностей рекомбинации (см. рис. 1-23). Одна из них начинается переходом электрона на ловушку и превращением ее в отрицательный ион, другая — переходом дырки (т. е. уходом электрона) и превращением ловушки в положительный ион. Сравнительная интенсивность той или иной последовательности зависит от соотношения концентраций n и p (т. е. от положения уровня Ферми) и от соотношения концентраций «занятых» и свободных ловушек (т. е. от положения уровня Φ_t).

Поскольку в условиях квазинейтральности скорости накопления электронов и дырок одинаковы, приравняем правые части уравнений (1-56) и выразим отсюда вероятность F_t . Далее подставим в любое из уравнений (1-56) найденное значение F_t , а также коэф-

¹ Эти неравенства особенно очевидны из сравнения вторых слагаемых в правых частях уравнений (1-56), где коэффициенты C_{nR} и C_{pR} близки по значению, а вероятность F_t не зависит от соотношения концентраций n и p . Значит, если n и p сильно различаются, то будут сильно различаться вторые — рекомбинационные — слагаемые. Так как скорости накопления электронов и дырок в условиях квазинейтральности одинаковы, то аналогичное различие будет иметь место и у первых — генерационных — слагаемых.

коэффициенты C_{ng} , C_{pg} из соотношений (1-57а). Тогда после некоторых преобразований *скорость накопления* избыточных носителей можно представить в виде¹

$$V = \frac{d(\Delta n)}{dt} = \frac{d(\Delta p)}{dt} = - \frac{C_n C_p (np - n_0 p_0)}{C_n (n + n_t) + C_p (p + p_t)}, \quad (1-59a)$$

где у коэффициентов C для простоты опущен индекс R .

Если заменить коэффициенты C_n и C_p на обратные величины — времена жизни τ_{n0} и τ_{p0} , то выражение (1-59а) будет иметь вид:

$$V = - \frac{np - n_0 p_0}{(n + n_t) \tau_{p0} + (p + p_t) \tau_{n0}}. \quad (1-59б)$$

Время жизни. Неравновесное время жизни найдем с помощью определения (1-54). Предварительно подставим в (1-59б) выражения (1-48) и учтем равенство (1-50а). Тогда

$$\tau = \frac{\Delta n}{-V} = \frac{n_0 + n_t + \Delta n}{n_0 + p_0 + \Delta n} \tau_{p0} + \frac{p_0 + p_t + \Delta n}{n_0 + p_0 + \Delta n} \tau_{n0}. \quad (1-60a)$$

Для малых возмущений ($\Delta n \rightarrow 0$) получим выражение, соответствующее (1-52):

$$\tau_0 = \frac{n_0 + n_t}{n_0 + p_0} \tau_{p0} + \frac{p_0 + p_t}{n_0 + p_0} \tau_{n0}. \quad (1-60б)$$

Для больших возмущений ($\Delta n \rightarrow \infty$) получаем:

$$\tau_\infty = \tau_{n0} + \tau_{p0}. \quad (1-60в)$$

Прежде чем анализировать время жизни избыточных носителей τ , остановимся на его составляющих τ_{n0} и τ_{p0} . Учитывая выражение для коэффициентов C_R в формулах (1-56), получаем:

$$\left. \begin{aligned} \tau_{n0} &= C_n^{-1} = \frac{1}{r_n N_t}; \\ \tau_{p0} &= C_p^{-1} = \frac{1}{r_p N_t}. \end{aligned} \right\} \quad (1-61)$$

Физический смысл параметров τ_{n0} и τ_{p0} можно понять из следующих рассуждений.

Пусть имеется электронный полупроводник, у которого $n_0 \gg \gg p_0$, n_t , p_t ; тогда из (1-60б) получается $\tau_0 \approx \tau_{p0}$. В случае дырочного полупроводника получается $\tau_0 \approx \tau_{n0}$. Следовательно, τ_{n0} и τ_{p0} — это времена жизни неосновных носителей в ярко выраженных примесных полупроводниках при малых возмущениях равновесия.

Аналогичный вывод получился для случая непосредственной рекомбинации, когда мы анализировали ее «вероятностным методом» [см. замечания к формулам (1-44)]. Однако существенная

¹ В литературе величину V часто называют *скоростью* или *темпом рекомбинации*, что может запутать неискующего читателя, так как скорость рекомбинации — лишь одна из составляющих скорости накопления. Значение $V < 0$ соответствует рассасыванию.

разница состоит в том, что при ловушечном механизме параметры τ_{n0} и τ_{p0} зависят не от плотности состояний в энергетических зонах, а от концентрации рекомбинационных центров. Из выражений (1-6) очевидно, что рост концентрации ловушек приводит к уменьшению времени жизни τ_{n0} и τ_{p0} , а следовательно, и результирующего времени жизни τ . Ясно также, что τ_{n0} и τ_{p0} слабо зависят от температуры, поскольку в первом приближении $r \sim v_t \sim \sqrt{T}$ [см. (1-27)].

В отличие от непосредственной рекомбинации коэффициенты r_n и r_p при ловушечном механизме имеют разные значения. Поэтому времена жизни τ_{n0} и τ_{p0} могут существенно различаться. Причиной такого различия являются неодинаковые сечения захвата и неодинаковые эффективные массы электронов и дырок [см. (1-27) и сноску¹ на с. 48].

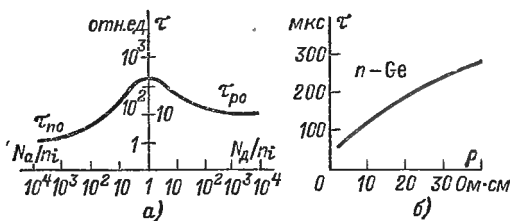


Рис. 1-24. Зависимость времени жизни от концентраций донорных и акцепторных примесей (а) и удельного сопротивления (б).

Выражения (1-60) позволяют оценить зависимость времени жизни избыточных носителей τ от таких важных факторов, как концентрация примесей (или удельное сопротивление материала), температура и концентрация избыточных носителей. Предположим для определенности, что уровни ловушек расположены в верхней половине запрещенной зоны на расстоянии нескольких φ_T от ее середины. Тогда из соотношений (1-58) получаем: $n_t \gg n_i$ и $p_t \ll \ll n_i$.

Рассмотрим зависимость $\tau(n_0)$ с помощью выражения (1-60б). Подставляя в него $p_0 = n_i^2/n_0$, $p_t = n_i^2/n_t$ и дифференцируя τ по n_0 , приходим к выводу, что функция $\tau(n_0)$ имеет максимум вблизи собственной концентрации, а именно при

$$n_0 \approx n_i \left(1 + \frac{\tau_{p0} - \tau_{n0}}{a\tau_{p0} + \tau_{n0}/a} \right),$$

где $a = n_t/n_i \gg 1$. Положение максимума зависит от соотношения времен жизни τ_{p0} и τ_{n0} .

Поскольку концентрации свободных носителей n_0 и p_0 являются однозначными и монотонными функциями концентраций соответствующих примесей [см. (1-24) для доноров], зависимость $\tau(N)$ (рис. 1-24, а) оказывается аналогичной зависимости $\tau(n_0)$. На рис. 1-24, б для примера приведена кривая $\tau(\rho)$ для электронного германия.

Физическая причина зависимости $\tau(N)$ состоит в следующем. В электронном полупроводнике по мере роста концентрации доноров электроны занимают все большую долю ловушечных уровней. При этом облегчаются захват дырок и их рекомбинация с электронами — последовательность, показанная на рис. 1-23 справа. Результирующее время жизни уменьшается, поскольку оно определяется рекомбинацией не основных носителей [см. замечания к (1-52)]. В дырочном полупроводнике рост концентрации акцепторов сопровождается уменьшением доли занятых ловушек. Соответственно облегчаются захват и рекомбинация электронов (последовательность, показанная на рис. 1-23 слева) и уменьшается время жизни.

Зависимость времени жизни от температуры также следует из (1-60б). Так, у электронного полупроводника, у которого при докритических температурах (см. § 1-7) $n_0 = N_d = \text{const}$, зависимость $\tau(T)$ обусловлена функцией $n_t(T)$. Согласно (1-57б) с ростом

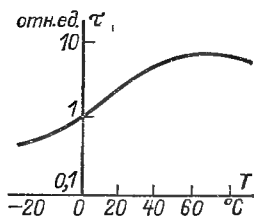


Рис. 1-25. Зависимость времени жизни от температуры.

температуры концентрация n_t растет; следовательно, будет увеличиваться и время жизни. В области низких температур, когда $n_t > n_0$, зависимость $\tau(T)$ проявляется слабо. Она становится существенной при температурах, соответствующих условию $n_t \approx n_0$ или, что то же самое, $\varphi_F \approx \varphi_t$. При еще более высоких температурах, когда $n_t \gg n_0$, время жизни возрастает столь же сильно (экспоненциально), как и концентрация n_t (рис. 1-25). В интервале $\pm 60^\circ\text{C}$ изменения времени жизни могут достигать 1—2 порядков.

В области сверхкритических температур концентрация n_0 начинает увеличиваться [см. (1-24а)]; соответственно рост функции $\tau(T)$ замедляется, а затем переходит в спад. Аналогичные выводы относятся и к дырочным полупроводникам.

Из сказанного следует, что в полупроводниках с большей концентрацией примеси, т. е. с меньшим удельным сопротивлением, зависимость времени жизни от температуры сдвигается в область более высоких температур, а значит, в рабочем диапазоне она менее существенна.

Физическая причина зависимости $\tau(T)$ состоит в следующем. В электронном полупроводнике с ростом температуры фононы все больше деионизируют ловушки, «срывая» с них электроны. Доля занятых электронами ловушек уменьшается, затрудняется рекомбинация дырок и результирующее время жизни растет [см. для сравнения причину зависимости $\tau(N)$]. В дырочном полупроводнике с ростом температуры фононы способствуют за пол н е н и ю ловушек, переводя на них электроны из валентной зоны. Тем самым затрудняется рекомбинация свободных электронов и время жизни увеличивается.

Наконец, рассмотрим зависимость времени жизни от концентрации избыточных носителей Δn . Из выражений (1-60а) и (1-60б) можно получить соотношение

$$\tau = \tau_0 \frac{1 + a \Delta n}{1 + b \Delta n} \quad (1-62)$$

Следовательно,

$$\tau_{\infty} = \tau_0 \frac{a}{b}, \quad (1-63)$$

где

$$a = \frac{\tau_{n0} + \tau_{p0}}{\tau_{p0}(n_0 + n_i) + \tau_{n0}(p_0 + p_i)}; \quad b = \frac{1}{n_0 + p_0}.$$

Как видим, возможны три случая: $\tau_0 > \tau_{\infty}$ (если $a < b$); $\tau_0 < \tau_{\infty}$ (если $a > b$) и $\tau_0 = \tau_{\infty}$ (если $a = b$). В ярко выраженных примесных полупроводниках время жизни возрастает с ростом Δn , поскольку в таких полупроводниках $\tau_0 = \tau_{n0}$ или $\tau_0 = \tau_{p0}$, тогда как $\tau_{\infty} = \tau_{p0} + \tau_{n0}$. С уменьшением концентрации примесей (т. е. с повышением удельного сопротивления) эта зависимость ослабляется, а затем переходит в обратную — время жизни уменьшается с ростом Δn . Кривые на рис. 1-26, а иллюстрируют зависимость τ_{∞}/τ_0 от удельного сопротивления ρ [19].

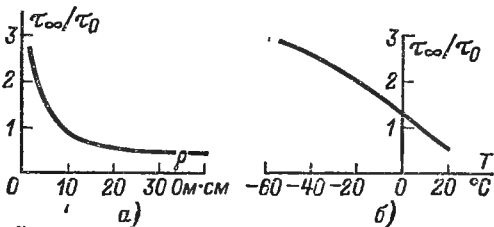


Рис. 1-26. Отношение времени жизни при высоком уровне инжекции ко времени жизни при низком уровне инжекции как функция удельного сопротивления (а) и температуры (б).

Поскольку τ_{∞} (как и τ_{n0} , τ_{p0}) слабо зависит от температуры, а τ_0 резко возрастает с температурой, то соотношение τ_{∞}/τ_0 получается существенно различным при разных температурах для одного и того же материала (рис. 1-26, б).

Все рассмотренные зависимости приходится учитывать при анализе и расчете как полупроводниковых приборов, так и электронных схем с их использованием.

Поверхностная рекомбинация. Законы генерации — рекомбинации на поверхности кристалла в основном такие же, как и в его объеме. В частности, скорость накопления носителей (на единицу площади) выражается формулой [20]

$$V = - \frac{C_n C_p (pn - p_0 n_0)}{C_n (n_{s0} + n_i) + C_p (p_{s0} + p_i)}, \quad (1-64)$$

близкой к формуле (1-59а)*. Однако внешние проявления процессов генерации — рекомбинации на поверхности и их количественная характеристика имеют некоторые особенности и заслуживают специального рассмотрения.

* Отличия (1-64) от (1-59а) состоят в том, что скорость накопления в (1-64) имеет размерность $\text{см}^{-2} \cdot \text{с}^{-1}$ (т. е. относится к единице площади), а коэффициенты C_n и C_p имеют размерность $\text{см}/\text{с}$ (как и скорость поверхностной рекомбинации, см. с. 46). Концентрации, стоящие в знаменателе (1-64), относятся к тонкому приповерхностному слою (индекс s), а концентрации, стоящие в числителе, — к объему полупроводника.

Дело в том, что поверхность полупроводника (как особая дефектная область) представляет собой чрезвычайно тонкий слой толщиной в несколько межатомных расстояний, т. е. несколько ангстрем. Поэтому любое реальное нарушение равновесия (под действием освещения и т. п.) происходит не на «самой» поверхности, а на некотором расстоянии от нее — в приповерхностном объеме кристалла. В результате между объемом (где равновесие нарушено и имеются избыточные носители) и поверхностью возникает поток частиц в сторону их меньшей концентрации¹. Такой поток удобно характеризовать скоростью поверхностной рекомбинации S , определяемой как поток частиц (на единицу площади поверхности и в единицу времени), отнесенный к приращению концентрации в приповерхностном объеме². Подставляя (1-48) в (1-64) и учитывая (1-50а), получаем:

$$S = \frac{-V}{\Delta n} = C_n \frac{C_p (n_0 + p_0 + \Delta n)}{C_n (n_{s0} + n_t) + C_p (p_{s0} + p_t)}. \quad (1-65)$$

Скорость поверхностной рекомбинации как параметр удобна тогда, когда нужно специально вычислить ток, обусловленный движением носителей к поверхности. Такой ток очень просто связан с параметром S :

$$I_s = qA |V| = qAS \Delta n,$$

где A — площадь, на которой имеет место рекомбинация.

Однако во многих случаях важно решить другую задачу: оценить результирующую скорость накопления (или рассасывания) носителей в некотором объеме, когда эта скорость частично обусловлена рекомбинацией в данном объеме, а частично — уходом носителей из данного объема к поверхности. В таких случаях, как уже отмечалось, удобно пользоваться эффективным временем жизни (1-38). Этот параметр получается из очевидного соотношения

$$\frac{d(\Delta n)}{dt} = \left(\frac{d(\Delta n)}{dt} \right)_v + \left(\frac{d(\Delta n)}{dt} \right)_s,$$

если записать его в виде

$$\frac{\Delta n}{\tau} = \frac{\Delta n}{\tau_v} + \frac{\Delta n}{\tau_s},$$

¹ Разумеется, часть потока направлена и в глубь кристалла. Если последний имеет достаточную толщину, то поверхностная рекомбинация будет сочетаться с объемной, т. е. будет не столь ярко выражена, как в тонкой пластинке, где поток неизбежно попадает на обе противоположащие поверхности.

² Последняя оговорка не случайна: если нарушение равновесия происходит вдали от поверхности [практически на расстоянии, большем диффузионной длины (1-11)], то поверхностные явления оказываются несущественными и можно считать $S = 0$. С физической точки зрения это значит, что избыточная концентрация рассасывается (благодаря объемной рекомбинации) до того, как носители достигнут поверхности.

где индексы v и s соответствуют величинам, связанным с объемом и поверхностью.

В общем виде установить связь между скоростью поверхностной рекомбинации S и поверхностным временем жизни τ_s трудно [21]. Эта задача решается сравнительно просто лишь для двух частных, но важных вариантов геометрии образца, к которым и стараются свести реальные варианты при практических расчетах полупроводниковых приборов.

Первый вариант — брусок, у которого размеры по осям x и y конечны и равны a и b , а протяженность по оси z бесконечная. Неравновесная концентрация носителей внутри кристалла считается одинаковой по всей оси z . Тогда, как показано в работе [2],

$$\frac{1}{\tau_s} = 4D \left(\frac{\eta_x^2}{a^2} + \frac{\eta_y^2}{b^2} \right), \quad (1-66)$$

где D — коэффициент диффузии носителей (см. § 1-11), а величины η определяются трансцендентными уравнениями:

$$\left. \begin{aligned} \eta_x \operatorname{tg} \eta_x &= \frac{Sa}{2D}; \\ \eta_y \operatorname{tg} \eta_y &= \frac{Sb}{2D}. \end{aligned} \right\} \quad (1-67)$$

Если скорость поверхностной рекомбинации S или размеры a и b достаточно малы, так что $Sa/2D < 0,5$ и $Sb/2D < 0,5$, то можно считать $\operatorname{tg} \eta = \eta$ и тогда

$$\frac{1}{\tau_s} \approx 2S \left(\frac{1}{a} + \frac{1}{b} \right). \quad (1-68a)$$

Если, наоборот, значения S , a и b достаточно велики, так что $Sa/2D > 5$ и $Sb/2D > 5$, то можно считать $\eta = \pi/2$ и тогда

$$\frac{1}{\tau_s} \approx \pi^2 D \left(\frac{1}{a^2} + \frac{1}{b^2} \right). \quad (1-68b)$$

В последнем случае поверхностное время жизни не зависит от скорости поверхностной рекомбинации и определяется только размерами и материалом образца.

Второй вариант — тонкая пластинка, у которой толщина по оси x равна a , а размеры по осям y и z бесконечны. Этот вариант можно рассматривать как частный случай бруска при $b = \infty$ или же анализировать самостоятельно [17]. В результате в общем виде получается:

$$\frac{1}{\tau_s} = \frac{4\eta_x D}{a^2}, \quad (1-69)$$

а в частных случаях

$$\frac{1}{\tau_s} \approx \frac{2S}{a}, \quad \text{если } \frac{Sa}{2D} \leq 0,5 \quad (1-70a)$$

или

$$\frac{1}{\tau_s} \approx \frac{\pi^2 D}{a^2}, \quad \text{если } \frac{Sa}{2D} \geq 5. \quad (1-70b)$$

Формулами (1-69) и (1-70) часто пользуются для расчета значений S по измеренным значениям τ и τ_v (из которых легко получить τ_s) и известным a и D :

$$S = \sqrt{\frac{D}{\tau_s}} \operatorname{tg} \frac{a}{2\sqrt{D\tau_s}}.$$

Если, например, $a = 0,01$ см; $D = 90$ см²/с; $\tau_p = 100$ мкс и $\tau = 10$ мкс, то $\tau_s \approx 10$ мкс и $S \approx 500$ см/с.

На рис. 1-27 приведены зависимости S от некоторых физических параметров. Эти зависимости можно объяснить, руководствуясь теми же соображениями, что и при анализе аналогичных зависимостей объемного времени жизни в предыдущем разделе. «Обратный» характер функций $S(N)$ и $S(T)$ по сравнению с функциями $\tau(N)$ и $\tau(T)$ обусловлен тем, что по определению скорость поверхностной рекомбинации прямо пропорциональна, тогда как время жизни обратно пропорционально скорости накопления V [см. (1-60а) и (1-65)].

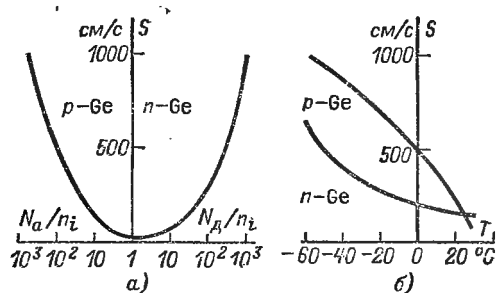


Рис. 1-27. Зависимость скорости поверхностной рекомбинации от концентрации донорных и акцепторных примесей (а) и от температуры (б).

Естественно, что скорость поверхностной рекомбинации существенно зависит от способа и качества обработки поверхности кристалла.

1-11. ЗАКОНЫ ДВИЖЕНИЯ НОСИТЕЛЕЙ ЗАРЯДА В ПОЛУПРОВОДНИКАХ

В общем случае движение носителей заряда в полупроводниках обусловлено двумя процессами: *диффузией* под действием градиента концентрации и *дрейфом* под действием градиента потенциала в электрическом поле. Поскольку в полупроводниках мы имеем дело с двумя типами носителей — дырками и электронами, полный ток состоит из четырех составляющих ¹:

$$i = (j_p)_{\text{диф}} + (j_p)_{\text{др}} + (j_n)_{\text{диф}} + (j_n)_{\text{др}}, \quad (1-71)$$

где индексы «диф» и «др» относятся соответственно к диффузионной и дрейфовой составляющим токов.

Плотности дрейфовых составляющих тока пропорциональны градиенту электрического потенциала ϕ , т. е. напряженности электрического поля E . В одномерном случае, когда движение носителей происходит только вдоль оси x , без отклонения в стороны, имеем:

$$(j_p)_{\text{др}} = -q\mu_p p \frac{\partial \phi}{\partial x} = q\mu_p p E; \quad (1-72a)$$

$$(j_n)_{\text{др}} = -q\mu_n n \frac{\partial \phi}{\partial x} = q\mu_n n E. \quad (1-72б)$$

¹ При анализе обычно удобнее пользоваться не токами i , а плотностями токов j , что и сделано в формуле (1-71). Там, где это не вызывает недоразумений, будем для краткости называть величину j током.

Следует заметить, что в обычном соотношении $E = -d\varphi/dx$, использованном в формулах (1-72), потенциал φ относится к положительным зарядам, тогда как на зонных диаграммах величина φ характеризует энергию отрицательных зарядов — электронов. Поэтому для зонных диаграмм действительно соотношение $E = d\varphi/dx$, т. е. положительной напряженности поля соответствует положительный градиент энергетических уровней.

Плотности диффузионных составляющих тока пропорциональны градиентам химических потенциалов χ_n и χ_p , которые для невырожденных полупроводников выражаются формулами (1-12). Поэтому в одномерном случае имеем¹:

$$(j_p)_{\text{диф}} = -q\mu_p \frac{d\chi_p}{dx} = -qD_p \frac{dp}{dx}; \quad (1-73a)$$

$$(j_n)_{\text{диф}} = q\mu_n \frac{d\chi_n}{dx} = qD_n \frac{dn}{dx}. \quad (1-73б)$$

Здесь D_p, D_n — коэффициенты диффузии дырок и электронов, связанные с подвижностями тех же носителей формулой Эйнштейна

$$D = \varphi_T \mu. \quad (1-74)$$

Зависимость коэффициентов диффузии от температуры легко получается с помощью (1-3) и (1-32):

$$D = D_0 \left(\frac{T_0}{T} \right)^{1/2}. \quad (1-75a)$$

Зависимость коэффициентов диффузии от концентрации примесей определяется формулой (1-33a), если заменить μ на D :

$$D = D_0 \left(\frac{N_0}{N} \right)^{1/3}. \quad (1-75б)$$

Из выражений (1-73) следует, что в невырожденных полупроводниках, для которых действительно использованные значения химических потенциалов, диффузионные токи пропорциональны градиенту концентраций носителей, а коэффициенты диффузии не зависят от этих концентраций.

Подставляя (1-72) и (1-73) в формулу (1-71), получаем плотность полного тока:

$$j = -qD_p \frac{\partial p}{\partial x} + q\mu_p E + qD_n \frac{\partial n}{\partial x} + q\mu_n E. \quad (1-76)$$

Как видим, для определения тока необходимо знать концентрации носителей и напряженность поля. В общем случае концентрации p и n зависят от двух переменных: координаты x и времени t . Поэтому для определения токов нужно предварительно найти

¹ Знак минус в формуле (1-73a) имеет следующий физический смысл: диффузия всегда происходит в направлении убывания концентрации, а поскольку дырки несут положительный заряд, ток $(j_p)_{\text{диф}}$ должен быть положительным при $dp/dx < 0$.

функции $p(x, t)$ и $n(x, t)$. Эти функции являются решениями так называемых *уравнений непрерывности потока*, которым в любой момент времени подчиняется движение носителей.

Для дырок и электронов уравнения непрерывности записываются в следующем виде¹:

$$\frac{\partial p}{\partial t} = \Delta g_p - \frac{p-p_0}{\tau_p} - \frac{1}{q} \operatorname{div}(j_p); \quad (1-77a)$$

$$\frac{\partial n}{\partial t} = \Delta g_n - \frac{n-n_0}{\tau_n} + \frac{1}{q} \operatorname{div}(j_n), \quad (1-77б)$$

где $p-p_0 = \Delta p$ и $n-n_0 = \Delta n$ — избыточные концентрации [см. (1-48)]; Δg_p и Δg_n — скорости генерации под действием внешних факторов, например света [скорости генерации под действием внутренних факторов — фоонов — учтены членами p_0/τ_p и n_0/τ_n , см. (1-43)].

Слагаемые в правых частях (1-77) соответствуют возможным причинам изменения концентрации носителей во времени. В частности, последние слагаемые можно рассматривать как скорости накопления или рассасывания носителей, обусловленные неравенством потоков, втекающих и вытекающих из некоторого элементарного объема. Такой небаланс потоков характеризуется дивергенцией вектора плотности потока. В нашем случае плотность потока есть j/q . Дивергенция этого вектора в одномерном случае равна:

$$\operatorname{div} \frac{j}{q} = \frac{1}{q} \frac{\partial}{\partial x} [(j)_{\text{диф}} + (j)_{\text{др}}].$$

Подставляя сюда соотношения (1-72) и (1-73), получаем:

$$\frac{1}{q} \operatorname{div}(j_p) = -D_p \frac{\partial^2 p}{\partial x^2} + \mu_p E \frac{\partial p}{\partial x} + p \mu_p \frac{\partial E}{\partial x};$$

$$\frac{1}{q} \operatorname{div}(j_n) = D_n \frac{\partial^2 n}{\partial x^2} + \mu_n E \frac{\partial n}{\partial x} + n \mu_n \frac{\partial E}{\partial x}.$$

С учетом этих выражений, а также при отсутствии внешних факторов (свет, радиация и т. п.) уравнения непрерывности (1-77) принимают следующую форму:

$$\frac{\partial p}{\partial t} = -\frac{p-p_0}{\tau_p} + D_p \frac{\partial^2 p}{\partial x^2} - \mu_p E \frac{\partial p}{\partial x} - p \mu_p \frac{\partial E}{\partial x}; \quad (1-78a)$$

$$\frac{\partial n}{\partial t} = -\frac{n-n_0}{\tau_n} + D_n \frac{\partial^2 n}{\partial x^2} + \mu_n E \frac{\partial n}{\partial x} + n \mu_n \frac{\partial E}{\partial x}. \quad (1-78б)$$

В том случае, когда поле отсутствует или когда его ролью заведомо можно пренебречь, полагаем $E = 0$. При этом выражения

¹ Легко заметить, что уравнения непрерывности являются развитием уравнений накопления и рассасывания (1-47). Последние рассматривают явления только во времени для заданной точки, тогда как первые благодаря наличию члена $\operatorname{div} j$ рассматривают явления и во времени, и в пространстве и потому являются уравнениями в частных производных [см. (1-78) и (1-79)].

(1-78) упрощаются и носят название *уравнений диффузии*:

$$\frac{\partial p}{\partial t} = -\frac{p-p_0}{\tau_p} + D_p \frac{\partial^2 p}{\partial x^2}; \quad (1-79a)$$

$$\frac{\partial n}{\partial t} = -\frac{n-n_0}{\tau_n} + D_n \frac{\partial^2 n}{\partial x^2}. \quad (1-79b)$$

Уравнения (1-79) позволяют достаточно строго анализировать многие типы полупроводниковых приборов. В тех случаях, когда полем пренебречь нельзя, пользуются полными уравнениями (1-78). Если напряженность E меняется вдоль оси x (т. е. если в полупроводнике имеется существенный объемный заряд), приходится дополнительно привлекать *уравнение Пуассона*, которое в одномерном случае имеет вид:

$$\frac{dE}{dx} = \frac{\lambda}{\epsilon_0 \epsilon}, \quad (1-80)$$

где λ — плотность заряда¹; ϵ_0 — электрическая постоянная (табл. 1-2); ϵ — относительная диэлектрическая проницаемость.

В общем случае

$$\lambda = q(p + N_d^* - n - N_a^*). \quad (1-81a)$$

Поскольку в условиях нейтральности $\lambda = 0$, можно считать, что объемный заряд есть следствие приращения концентраций в правой части (1-81a). Если степень ионизации примесей неизменна (т. е. $N_{d,a}^* = \text{const}$), то

$$\lambda = q(\Delta p - \Delta n). \quad (1-81b)$$

Решение системы уравнений (1-78) и (1-80) в общем виде невозможно. В каждом конкретном случае приходится вводить те или иные упрощения. В следующих параграфах рассматривается несколько примеров таких упрощенных решений, которые одновременно позволят ознакомиться еще с некоторыми важными свойствами и параметрами полупроводников.

1-12. ОБЪЕМНЫЕ ЗАРЯДЫ И ПОЛЯ В ПОЛУПРОВОДНИКАХ

Диэлектрическая релаксация. Пусть в ограниченном объеме полупроводника удалось сосредоточить избыточные концентрации электронов и дырок, так что образовался объемный заряд с плотностью λ . Под действием возникшего поля заряд будет рассасываться, т. е. носители будут покидать тот начальный объем, в котором они были сосредоточены. Такое рассасывание заряда под действием собственного поля носит название *диэлектрической релаксации*, или релаксации Максвелла.

¹ В электродинамике плотность заряда обычно обозначают через ρ . В полупроводниковой технике такое обозначение нежелательно, так как оно принято для удельного сопротивления.

При анализе диэлектрической релаксации пренебрегают комбинацией носителей и их диффузией, чтобы выделить явление в чистом виде. Следовательно, в правых частях (1-78) можно опустить все члены, кроме последних:

$$\frac{\partial p}{\partial t} = -p\mu_p \frac{\partial E}{\partial x}; \quad \frac{\partial n}{\partial t} = n\mu_n \frac{\partial E}{\partial x}.$$

Вычитая второе уравнение из первого, подставляя dE/dx из (1-80), λ из (1-81б) и учитывая (1-35), получаем уравнение релаксации¹:

$$\frac{d(\Delta p - \Delta n)}{dt} = -\frac{\sigma}{\epsilon_0 \epsilon} (\Delta p - \Delta n). \quad (1-82)$$

Решением является экспоненциальная функция

$$\Delta p - \Delta n = [\Delta p(0) - \Delta n(0)] e^{-t/\tau_e}, \quad (1-83)$$

где $[\Delta p(0) - \Delta n(0)]$ — начальная избыточная концентрация и

$$\tau_e = \epsilon_0 \epsilon / \sigma \quad (1-84)$$

— время диэлектрической релаксации.

Величина τ_e характеризует время, в течение которого нарушена нейтральность полупроводника: через $(3 \div 4) \tau_e$ объемный заряд практически рассасывается и нейтральность восстанавливается.

Как видно из (1-84), время релаксации помимо диэлектрической проницаемости зависит от удельной проводимости или удельного сопротивления. Например, если $\rho = 1$ Ом·см, то для германия и кремния $\tau_e \approx 10^{-12}$ с. Такое крайне малое значение типично для процессов диэлектрической релаксации и является одной из основ квазинейтральности полупроводников (см. конец с. 31).

С ростом удельного сопротивления время релаксации увеличивается и у собственных полупроводников (особенно кремния) должно, казалось бы, достигать сравнительно больших значений. Следует, однако, иметь в виду, что наличие избыточных электронов и дырок в возмущенном объеме приводит к уменьшению удельного сопротивления на этом участке и, следовательно, согласно (1-84) способствует уменьшению времени релаксации. Поэтому даже в собственных полупроводниках при сколько-нибудь значительных возмущениях значение τ_e обычно имеет порядок 10^{-12} с.

Изменение удельного сопротивления полупроводника на том участке, где по тем или иным причинам скапливаются избыточные носители, носит название эффекта *модуляции проводимости*. Этот эффект играет значительную роль в полупроводниковых приборах, особенно при больших сигналах (см., например, § 2-8).

¹ Обычно в литературе структуру объемного заряда не расшифровывают, а записывают уравнение (1-82) в виде

$$\frac{\partial \rho}{\partial t} = -\operatorname{div} E = -\frac{\sigma}{\epsilon_0 \epsilon} \rho,$$

где ρ — плотность объемного заряда.

Из структуры выражения (1-83) следует важный вывод о двух вариантах механизма релаксации. А именно, предположим, что полупроводник дырочный и в нем создано начальное возмущение $\Delta p(0)$ при $\Delta n(0) = 0$. Тогда, поскольку электронов в окружающем объеме очень мало и они не могут нейтрализовать (скомпенсировать) возмущение $\Delta p(0)$, релаксация будет состоять в уменьшении концентрации Δp благодаря уходу избыточных дырок из начального объема. Процесс рассасывания будет описываться выражением

$$\Delta p = \Delta p(0) e^{-t/\tau_e}$$

и закончится через время, равное примерно $3\tau_e$ (рис. 1-28, а).

Предположим теперь, что в том же дырочном полупроводнике начальное возмущение представляет собой избыточную концентрацию электронов $\Delta n(0)$ при $\Delta p(0) = 0$. Тогда релаксация будет состоять в нейтрализации (компенсации) заряда электронов $\Delta n(0)$ дырками, приходящими под действием поля из внешнего объема, где их очень много. Процесс нейтрализации описывается выражением

$$\Delta p = \Delta n(0) (1 - e^{-t/\tau_e}),$$

т. е. характеризуется увеличением концентрации Δp от 0 до $\Delta n(0)$ и заканчивается, как и в первом случае, через $3\tau_e$. После этого с гораздо меньшей скоростью, определяемой временем жизни τ , будет происходить рекомбинация электронов и дырок, скопившихся в области возмущения¹ (рис. 1-28, б).

Таким образом, если начальное возмущение вызвано основными носителями, то процессы рассасывания объемного заряда и избыточной концентрации происходят с одной и той же малой постоянной времени τ_e . Если же начальное возмущение вызвано

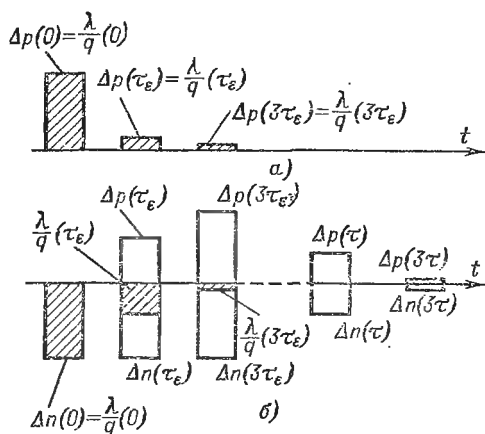


Рис. 1-28. Схематическое изображение процессов диэлектрической релаксации в примесном (дырочном) полупроводнике.

а — исходное возмущение вызвано основными носителями; б — исходное возмущение вызвано неосновными носителями.

¹ Заметим, что нейтрализация в принципе не может быть полной, иначе нельзя было бы объяснить наличие градиента концентрации дырок между рассматриваемым объемом и основной частью кристалла. В действительности остается небольшой нескомпенсированный заряд электронов, поле которого и поддерживает градиент концентрации дырок. Именно неполной компенсацией заряда объясняется термин *квазинейтральность*. Электрическое поле, обусловленное неполной компенсацией заряда, будем называть *остаточным*.

и основными носителями, то нейтрализация объемного заряда заканчивается столь же быстро, с малой постоянной времени τ_e , а рассасывание обеих избыточных концентраций — значительно медленнее (уже в условиях квазинейтральности) с постоянной времени τ . Второй вариант процесса является одной из основ транзисторной техники.

Начальное возмущение может выражаться не только в увеличении, но и в уменьшении соответствующих концентраций. При этом в уравнении (1-83) меняются только знаки при избыточных концентрациях; характер процессов остается прежним.

Введение («впрыскивание») новых носителей в примесный полупроводник называют *инжекцией*, а их «отсасывание» с образованием «отрицательного» возмущения — *экстракцией*.

Эффект поля. Пусть между металлической пластинкой и полупроводником задано напряжение U (рис. 1-29, а). Такая система представляет собой своеобразный конденсатор, у которого одна из обкладок полупроводниковая. На этой обкладке будет наведен такой же суммарный заряд, как и на металлической, но в отличие от металла заряд в полупроводнике не будет сосредоточен на поверхности, а распределится на некоторое расстояние в глубь кристалла. Распределение заряда, электрического поля и потенциала можно оценить, используя уравнение Пуассона (1-80). Специфика задачи в данном случае состоит в том, что плотность заряда λ не является заданной. Она согласно (1-81) определяется концентрациями подвижных носителей, а эти концентрации в свою очередь зависят от распределения потенциала φ_E . Поэтому прежде всего остановимся на зависимостях $n(\varphi_E)$ и $p(\varphi_E)$.

Заметим, что в рассматриваемой системе полупроводник, несмотря на наличие внутреннего индуцированного поля, находится в состоянии бальцмановского равновесия, так как нет результирующего тока. Значит, уровень Ферми во всех частях полупроводника постоянный: $\varphi_F = \text{const.}$ Для определения этого уровня воспользуемся формулами (1-18). Обозначим концентрации носителей и электростатический потенциал в правой, невозмущенной части полупроводника (рис. 1-29, а) через n_0 , p_0 и φ_{E0} и положим для простоты $\varphi_{E0} = 0$. Тогда из (1-18) находим

$$\varphi_F = \varphi_T \ln \frac{n_0}{n_i} = -\varphi_T \ln \frac{p_0}{n_i}.$$

Подставляя найденное значение φ_F в (1-17), получаем искомые зависимости $n(\varphi_E)$ и $p(\varphi_E)$:

$$\left. \begin{aligned} n &= n_0 e^{-\varphi_E / \varphi_T}; \\ p &= p_0 e^{\varphi_E / \varphi_T}; \end{aligned} \right\} \quad (1-85)$$

потенциал φ_E отсчитывается от середины запрещенной зоны в нейтральной части полупроводника. Зависимости (1-85) позволяют решать уравнение Пуассона.

Рассмотрим сначала собственный полупроводник. В этом случае в (1-81a) следует положить $N_d^* = N_a^* = 0$ и подставить концентрации (1-85), считая $n_0 = p_0 = n_i$. Левую часть (1-80)

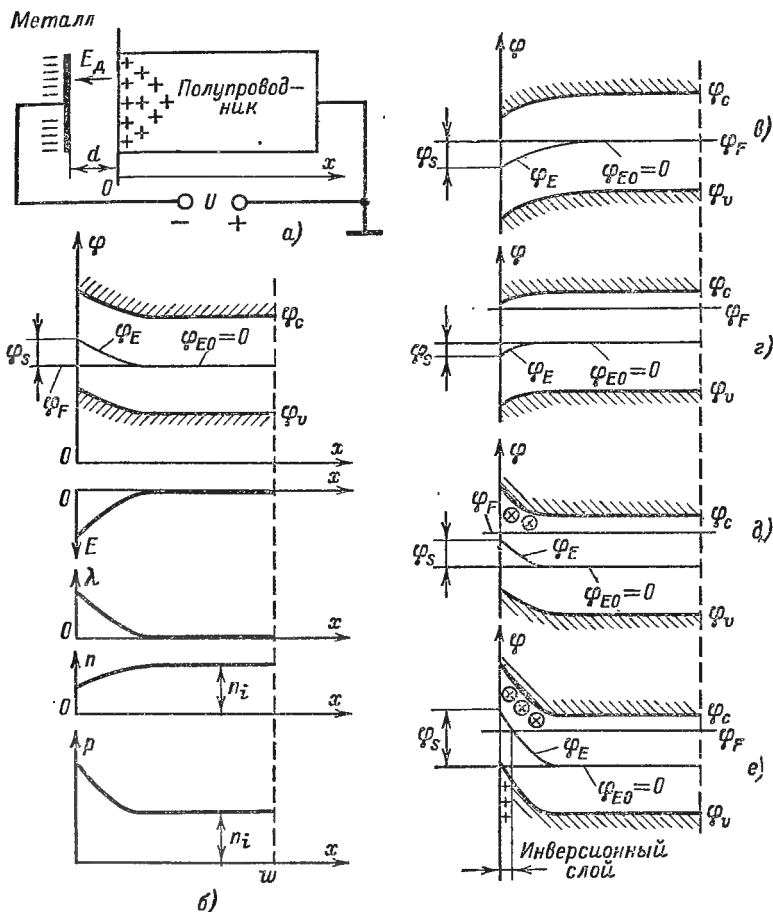


Рис. 1-29. Эффект поля в полупроводниках.

Схема эксперимента (a); распределение потенциала, поля, заряда и концентрации носителей в собственном полупроводнике при отрицательной (б) и положительной (в) полярностях напряжения на металлическом электроде; распределение потенциала в примесном (электронном) полупроводнике при положительной (г) и отрицательной (д, е) полярностях напряжения на металлическом электроде (модуль напряжения U для случая е значительно больше, чем для случая д).

запишем как $d^2\varphi_E/dx^2$, поделим обе части на φ_T и введем безразмерную переменную $\Phi = \varphi_E/\varphi_T$. Тогда уравнение Пуассона приводится к виду

$$\frac{d^2\Phi}{dx^2} = \frac{1}{2l_{Di}^2} (e^\Phi - e^{-\Phi}) = \frac{\text{Sh } \Phi}{l_{Di}^2}, \quad (1-86)$$

где

$$l_{Di} = \sqrt{\frac{\epsilon_0 \epsilon \varphi_T}{2qn_i}} \quad (1-87)$$

— *дебаевская длина* для собственного полупроводника.

Решение уравнения (1-86) несложно, но громоздко (см. [7]), поэтому ограничимся окончательным результатом:

$$\varphi_E(x) = 2\varphi_T \ln \frac{1 + \operatorname{th}(\varphi_s/4\varphi_T) e^{-x/l_{Di}}}{1 - \operatorname{th}(\varphi_s/4\varphi_T) e^{-x/l_{Di}}}, \quad (1-88a)$$

где $\varphi_s = \varphi_E(0)$ — так называемый *поверхностный потенциал* (рис. 1-29, б), т. е. напряжение между поверхностью кристалла (при $x = 0$) и его объемом (при $x = \infty$ или практически при $x > 3l_{Di}$).

Если выполняется условие $\varphi_s < 2\varphi_T$, то логарифм в (1-88a) можно разложить в ряд с точностью до членов 1-го порядка и получить элементарное выражение:

$$\varphi_E(x) \approx \varphi_s e^{-x/l_{Di}}. \quad (1-88б)$$

Из (1-88б) ясно видно, что дебаевская длина характеризует то расстояние, на котором потенциал φ_E еще существует, т. е. глубину проникновения электрического поля в полупроводник. Поэтому l_{Di} называют иногда *глубиной экранирования*, имея в виду, что слой объемного заряда «экранирует» основной объем полупроводника от проникновения внешнего электрического поля. Порядок собственной дебаевской длины для кремния: $l_{Di} \approx 20$ мкм, для германия: $l_{Di} \approx 0,6$ мкм.

Зная функцию $\varphi_E(x)$, легко получить функции $E(x)$, $\lambda(x)$, $p(x)$ и $n(x)$. Все они показаны на рис. 1-29, б. Если изменить полярность приложенного напряжения, то знак объемного заряда изменится и зоны искривятся в другую сторону (рис. 1-29, в).

Как видим, при любой полярности напряжения приповерхностный слой собственного полупроводника оказывается обогащенным подвижными носителями, т. е. имеет значительно большую удельную проводимость, чем основной объем. Изменение проводимости приповерхностного слоя под действием внешнего поля называют *эффектом поля*, а сам обогащенный слой — *каналом*. Поскольку канал характерен большой концентрацией электронов или дырок, можно сказать, что эффект поля в собственном полупроводнике приводит к образованию тонких искусственных слоев с проводимостью *n*- или *p*-типа. Такие искусственные слои играют важную роль в полупроводниковой технике. В частности, они находят применение в МДП транзисторах (см. гл.5).

Величину поверхностного потенциала φ_s можно оценить из условия

$$\epsilon_d E_d(0) = \epsilon_n E_n(0),$$

характеризующего непрерывность нормальной составляющей вектора электрической индукции на границе диэлектрика. Напряженность поля в диэлектрике

не зависит от x и определяется разностью потенциалов $U - \varphi_s$, поэтому

$$E_n(0) = \frac{U - \varphi_s}{d}.$$

Напряженность поля в полупроводнике определяется производной $d\varphi_E/dx$. Дифференцируя функцию (1-88a) и полагая $x = 0$, получаем:

$$E_n(0) = \frac{4\varphi_T}{l_{Di}} \frac{\text{th}(\varphi_s/4\varphi_T)}{1 - \text{th}^2(\varphi_s/4\varphi_T)}.$$

В результате поверхностный потенциал можно найти из трансцендентного уравнения

$$\frac{a}{2} \text{sh}(2z) + z = u, \quad (1-89a)$$

где

$$a = \frac{\epsilon_n d}{\epsilon_n l_{Di}}; \quad z = \frac{\varphi_s}{4\varphi_T}; \quad u = \frac{U}{4\varphi_T}.$$

Если параметр z достаточно мал ($z < 0,5$), то, считая $\text{sh}(2z) \approx 2z$, получаем:

$$\varphi_s \approx \frac{U}{a+1}. \quad (1-89б)$$

Если же параметры z и a достаточно велики ($z > 0,5$; $a > 2$), можно считать

$\text{sh}(2z) \approx \frac{1}{2} e^{2z}$ и $\frac{1}{2} a \text{sh}(2z) \gg z$; тогда

$$\varphi_s \approx 2\varphi_T \ln \frac{U}{a\varphi_T}. \quad (1-89в)$$

Сравнивая выражения (1-89б) и (1-89в), замечаем, что при малых внешних напряжениях поверхностный потенциал меняется почти пропорционально напряжению U , а затем, при напряжениях $U > 2(a+1)\varphi_T$, меняется весьма слабо. Оба эти вывода с физической точки зрения объясняются тем, что заряд металлической пластины всегда пропорционален напряжению U (если $U \gg \varphi_s$), тогда как равный ему заряд полупроводниковой пластины зависит от потенциала φ_s экспоненциально [см. (1-85)]. Значит, в области больших зарядов требуется гораздо меньшее изменение величины φ_s по сравнению с U , чтобы обеспечить одно и то же приращение заряда; получается зависимость типа (1-89в). В области малых зарядов, когда экспонента аппроксимируется прямой, необходимое изменение заряда обеспечивается почти одинаковыми изменениями значений φ_s и U ; получается зависимость типа (1-89б).

Заметим, что в рамках изложенной теории поверхностный потенциал, а значит, и искривление зон не должны превышать $\frac{1}{2}\varphi_s - (2-3)\varphi_T$. В противном случае уровень Ферми на поверхности (рис. 1-29) окажется вблизи или даже внутри одной из разрешенных зон. На этом участке полупроводник будет вырожденным и для него будут недействительными исходные предпосылки в виде формул (1-85).

Еще одно замечание касается параметра a [см. (1-89)]. Этот параметр существенно зависит от толщины диэлектрика d . Значение d не может быть произвольно малым: при условии $d < 0,01 \div 0,02$ мкм диэлектрик становится проницаемым для подвижных носителей благодаря туннельному эффекту (см. § 2-4).

При этом рассматриваемая структура перестает быть аналогом конденсатора и весь проведенный анализ теряет силу: обмен носителями через диэлектрик вызывает уменьшение накопленных зарядов вплоть до полного их исчезновения при $d \rightarrow 0$, т. е. при соприкосновении металлической обкладки с полупроводником, когда в системе протекает обычный ток проводимости.

Наконец, подчеркнем, что весь предыдущий анализ был основан на предположении, что искривление энергетических уровней в полупроводнике обусловлено только влиянием внешнего поля. На практике по ряду причин (см. § 2-4 и 5-5) искривление уровней имеет место и в отсутствие внешнего напряжения, т. е. в равновесном состоянии. Соответствующий поверхностный потенциал φ_{s0} называют *равновесным*.

Под действием внешнего напряжения этот начальный потенциал изменяется в ту или иную сторону. В частности, можно так подобрать полярность и значение напряжения, чтобы «скомпенсировать» потенциал φ_{s0} , т. е. сделать $\varphi_s = 0$ и тем самым ликвидировать искривление уровней. Такое напряжение называют *напряжением спрямления зон* U_F (англ. U_{FB} , где индекс — от слов Flat Band — плоские зоны).

Важнейшей особенностью эффекта поля в примесных полупроводниках по сравнению с собственными является возможность двух режимов — *обогащения и обеднения* приповерхностного слоя носителями.

Режим обогащения соответствует такой полярности приложенного напряжения, при которой основные носители притягиваются к поверхности и образуют тонкий канал с тем же типом проводимости, но с гораздо меньшим удельным сопротивлением. Этот случай близок к рассмотренному на рис. 1-29, б, в, но отличается от него меньшим искривлением зон, поскольку в примесном полупроводнике даже небольшие поверхностные потенциалы обеспечивают существенное сближение разрешенной зоны с уровнем Ферми, а значит, существенное увеличение концентрации основных носителей (рис. 1-29, г). Анализ показывает [7], что в случае малого искривления зон ($\varphi_s < 2\varphi_T$) потенциал вблизи поверхности примесного полупроводника описывается выражением (1-88б), но дебаевская длина l_{Di} заменяется на

$$l_D = \sqrt{\frac{\epsilon_0 \epsilon \varphi_T}{qn_0}} \quad (1-90a)$$

(для дырочного полупроводника концентрацию n_0 следует заменить на p_0).

Поскольку обычно $n_0 \gg n_i$, то дебаевская длина в примесных полупроводниках значительно меньше, чем в собственных. Например, при $n_0 = 10^{16} \text{ см}^{-3}$ из (1-90a) и для кремния, и для германия получаем $l_D \approx 0,04 \text{ мкм}$. Как видим, в обычных полупроводниках поле проникает на ничтожную глубину (0,1 мкм и менее).

Режим обеднения не соответствует обратной полярности напряжения, при которой основные носители отталкиваются от поверхности. В этом случае вместо тонкого канала с повышенной проводимостью образуется сравнительно толстый слой с пониженной проводимостью, в котором объемный заряд обусловлен ионами («обнаженными» атомами) примеси, от которых удалены основные носители (рис. 1-29, д). Протяженность этого слоя определяется не дебаевской длиной, а другим характеристическим параметром [ср. с (2-96)]:

$$l_0 = \sqrt{\frac{2\epsilon_0\epsilon\varphi_s}{qN_d}} \quad (1-90б)$$

(для дырочного полупроводника следует заменить N_d на N_a).

Хотя структура выражений (1-90а) и (1-90б) одинакова (тем более, что $n_0 \approx N_d$), но есть и существенная разница, состоящая в том, что дебаевская длина l_D зависит только от свойств полупроводника, тогда как толщина объемного заряда l_0 зависит еще и от приложенного напряжения, которое влияет на значение φ_s .

С ростом обратного напряжения объемный заряд примесных ионов растет, а заряд основных носителей продолжает уменьшаться, но одновременно растет заряд неосновных носителей, притягиваемых к поверхности. Когда нарастающий заряд неосновных носителей превысит заряд оставшихся основных, изменится тип проводимости приповерхностного слоя. Этот случай характеризуют термином *инверсия типа проводимости*, а канал, образованный неосновными носителями, называют *инверсионным слоем* (рис. 1-29, е). Протяженность инверсионного слоя обычно составляет всего 0,001—0,002 мкм.

У металлов с их огромными концентрациями свободных носителей (обычно 10^{22} см⁻³ и больше) характеристические длины l_D и l_0 должны были бы составлять доли ангстрема. В действительности, если учитывать квантовые эффекты, проявляющиеся при таких расстояниях, заряд сосредоточивается на значительно большей глубине, равной нескольким ангстремам, т. е. нескольким межатомным расстояниям.

Неоднородные полупроводники. В конце § 1-7 отмечалось, что в неоднородных полупроводниках должны иметь место внутренние электрические поля (см. рис. 1-17, а). Определим зависимость внутреннего поля от степени неоднородности на примере электронного полупроводника, у которого концентрация доноров выражается функцией $N_d(x)$. Концентрация электронов в общем случае будет выражаться несколько иной функцией, однако для простоты примем $n(x) = N_d(x)$.

В основу анализа положим больцмановское равновесие, т. е. постоянство уровня Ферми: $\varphi_F = \text{const}$. Дифференцируя по x обе части выражения (1-18а) и учитывая, что $d\varphi_E/dx = E$, получаем напряженность внутреннего поля в следующем виде:

$$E = -\varphi_T \frac{dn/dx}{n}. \quad (1-91)$$

Соответственно с помощью (1-80) находим плотность объемного заряда:

$$\lambda = \varepsilon_0 \varepsilon \varphi_T \left[\left(\frac{dn/dx}{n} \right)^2 - \frac{d^2 n/dx^2}{n} \right]. \quad (1-92)$$

Из выражений (1-91) и (1-92) видно, что внутреннее поле и объемный заряд тем больше, чем резче меняется концентрация основных носителей и чем меньше эта концентрация в каждой данной точке.

Для того чтобы получить поле с заданной напряженностью и заданной конфигурации, необходимо обеспечить определенное распределение концентрации примесей $N(x)$. В общем виде это еще не решенная технологическая задача. Практически приходится, наоборот, исходить из технологически реализуемых функций $N(x)$ и анализировать соответствующие им поля.

Рассмотрим случай

$$N_d(x) = N_d(0) e^{-x/L_d}, \quad (1-93)$$

который реализуется при диффузионной технологии (см. § 4-13). В этом случае, считая $n(x) = N_d(x)$, из (1-91) и (1-92) получаем:

$$E = \varphi_T / L_d; \quad (1-94)$$

$$\lambda = 0. \quad (1-95)$$

Как видим, при экспоненциальном распределении примеси электрическое поле имеет постоянную напряженность, а объемный заряд соответственно отсутствует. Поле будет тем сильнее, чем меньше длина L_d , характеризующая глубину проникновения примеси в кристалл. Для типичных значений $L_d = 1 \div 2$ мкм получается $E \approx 100 \div 200$ В/см.

Постоянное поле обусловлено, как и в обычном плоском конденсаторе, зарядами, сосредоточенными в тонких приповерхностных слоях при $x \approx 0$ и $x \approx w$ (рис. 1-30, а). Эти заряды образовались в результате небольшого сдвига электронов вдоль оси x под действием градиента концентрации dn/dx . Вблизи $x = 0$ остался нескомпенсированный (или не полностью скомпенсированный) положительный заряд «обнаженных» доноров, а вблизи $x = w$ получился отрицательный заряд накопившихся избыточных электронов. Толщины заряженных слоев определяются соответствующими характеристическими длинами¹. Эти длины различны при $x \approx 0$ и $x \approx w$ из-за различия концентраций $n(0)$ и $n(w)$, см. (1-93). На протяжении этих длин напряженность поля, постоянная в основной части кристалла, падает до нуля в точках $x = 0$ и $x = w$. Соответственно уровни φ_E , φ_c и φ_w , которые меняются линейно в основной части кристалла, в слоях объемного заряда претерпевают изгибы.

Структура заряженных слоев представляет первостепенный интерес при анализе полупроводников со ступенчатой неоднородностью. Пусть в точке $x = 0$ концентрация доноров скачком меняется от N_{d1} до N_{d2} (рис. 1-30, б) и пусть про-

¹ Параметры l_D и l_b (1-90) характеризуют не только экранирование полупроводника от внешнего поля, но и экранирование внешнего пространства от внутреннего поля в полупроводнике (см. рис. 1-30, а, где $E = 0$ слева от точки 0 и справа от точки w).

тяжести обоих однородных слоев слева и справа от этой точки очень велики (много больше соответствующих характеристических длин). Тогда в удаленных от $x = 0$ областях сохранится равновесное состояние, свойственное каждому из слоев, т. е. концентрации электронов будут соответственно n_{01} и n_{02} . Однако вблизи точки $x = 0$, где имеется большой (в принципе бесконечно большой) градиент

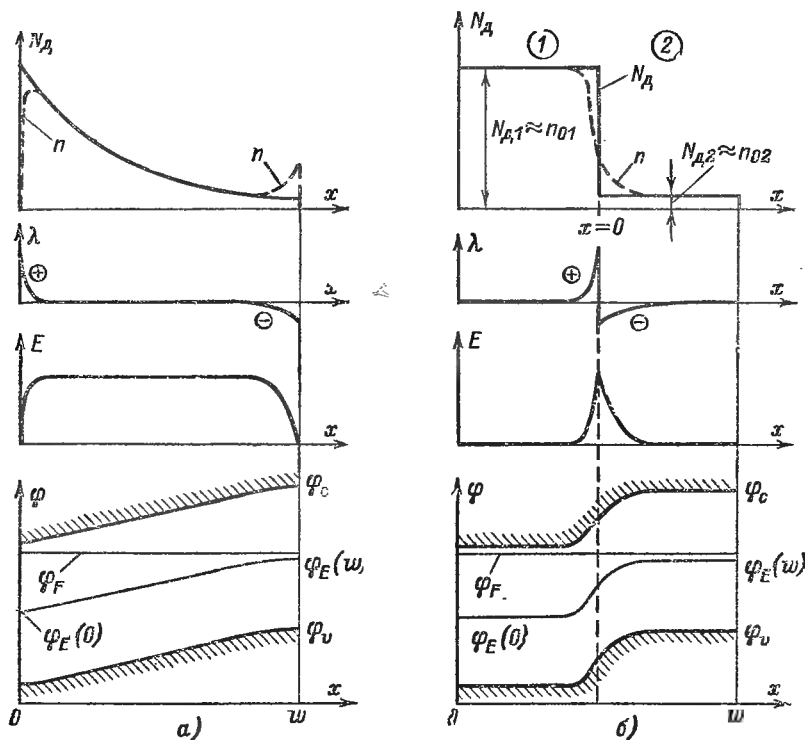


Рис. 1-30. Распределение примеси, заряда, поля и потенциала в неоднородных полупроводниках.

а — плавная (экспоненциальная) неоднородность; *б* — ступенчатая неоднородность.

концентрации электронов, состояние не может остаться неизменным: часть электронов перейдет из слоя 1 в слой 2. Следовательно, слева от точки $x = 0$ останется некоторый нескомпенсированный заряд доноров, а справа появится заряд избыточных электронов. Образуется двойной электрический слой толщиной, примерно равной сумме характеристических длин, в пределах которого действует электрическое поле и существует разность потенциалов (рис. 1-30, б). Сравнительно резкие изменения потенциала, обусловленные объемными зарядами, накладывают *потенциальными барьерами*.

В неоднородных полупроводниках характеристические длины l_D и l_0 не являются однозначными параметрами материала, поскольку концентрации n и N_d зависят от координат. Если, например, вычислить дебаевскую длину из выражения (1-90а), подставив значение $n(x)$, то она будет определять истинную протяженность объемного

заряда лишь в зависимости от изменения концентрации n на участке l_D невелико по сравнению с самой концентрацией:

$$\left| \frac{dn}{dx} \right| l_D < n. \quad (1-96a)$$

Если учесть (1-90a), то неравенство примет вид:

$$\left| \frac{dn}{dx} \right| < \sqrt{\frac{q}{\varepsilon_0 e \Phi_T}} n^{3/2} \approx 2 \cdot 10^{-3} n^{3/2} \quad (1-96б)$$

(численный коэффициент соответствует значениям $\varepsilon \approx 14$ и $T \approx 300$ K).

Чем сильнее выполняются неравенства (1-96), тем ярче проявляются экранирующие свойства полупроводника, т. е. обеспечивается локальная концентрация поля и заряда. Если же неравенства (1-96) выполняются слабо или нарушаются, то распределение потенциала делается весьма плавным, а поле и заряд все больше распространяются в глубь полупроводника, нарушая его квазинейтральность. В этом случае понятие характеристической длины вообще теряет определенность.

Применительно к распределению (1-93) условие (1-96a) запишется следующим образом:

$$L_d > l_D, \quad (1-97)$$

где значение l_D следует вычислять для участка с минимальной концентрацией примеси.

Квазинейтральность. Исходя из материала, изложенного в предыдущих разделах, можно сформулировать условие квазинейтральности, о котором уже не раз упоминалось.

Для однородных полупроводников условие квазинейтральности формулируется следующим образом:

В однородном полупроводнике существенные объемные заряды не могут иметь места в течение времени, большего нескольких времен диэлектрической релаксации, за исключением узких граничных участков протяженностью в несколько характеристических длин.

Можно сделать вывод: в однородном полупроводнике (с учетом сделанных оговорок) напряженность электрического поля может быть либо близкой к нулю, либо почти постоянной¹.

Сущ е с т в е н н о с т ь объемных зарядов, оговоренная в формулировке, с математической точки зрения означает необходимость учитывать члены с $\partial E / \partial x$ в уравнениях непрерывности (1-78). Следовательно, при анализе основного объема однородного полупроводника (вдали от границ) и в разумном интервале времени (не включающем начальный релаксационный интервал) членами с $\partial E / \partial x$ можно пренебрегать.

¹ Второй случай реализуется при подключении внешнего напряжения: как отмечалось в связи с рис. 1-17, б, протекание тока в однородном полупроводнике не вызывает перераспределения концентраций носителей.

Принцип квазинейтральности применительно к неоднородным полупроводникам должен, естественно, включать некоторое ограничение на степень неоднородности, поскольку согласно (1-92) в таких полупроводниках в общем случае могут иметь место значительные объемные заряды. Иначе говоря, нужно определить понятие с у щ е с т в е н н о г о объемного заряда. Логично предположить, что объемный заряд будет несущественным, если он значительно меньше того заряда, который соответствует полному количеству основных носителей в данной точке¹. Тогда правую часть выражения (1-92) следует положить много меньше qn , откуда условие квазинейтральности получится в виде

$$\frac{(dn/dx)^2 - n(d^2n/dx^2)}{n^3} \ll \frac{q}{\epsilon_0 \epsilon \varphi_T}. \quad (1-98)$$

Для всех реальных распределений $n(x)$ вторая производная d^2n/dx^2 положительна. Поэтому можно опустить второе слагаемое в левой части (1-98), усилив тем самым неравенство. Тогда условие квазинейтральности принимает форму

$$\left| \frac{dn}{dx} \right| \ll \sqrt{\frac{q}{\epsilon_0 \epsilon \varphi_T}} n^{3/2}, \quad (1-99)$$

совпадающую с формой условия применимости дебаевской длины (1-96б).

В итоге условие квазинейтральности для неоднородного полупроводника формулируется следующим образом:

Неоднородный полупроводник квазинейтрален в том же смысле, что и однородный, если степень его неоднородности соответствует неравенству (1-99).

1-13. КИНЕТИКА НОСИТЕЛЕЙ ЗАРЯДА В ПОЛУПРОВОДНИКАХ

Биполярная диффузия. Пусть на поверхность полупроводника падает рассеянный пучок света (рис. 1-31). Тогда в тонком приповерхностном слое, в который проникает свет, будут генерироваться электронно-дырочные пары со скоростью Δg , пар/(см³·с). Между поверхностью и основным объемом возникнут градиенты концентрации электронов и дырок и избыточные носители начнут диффундировать в глубь полупроводника. Такое с о в м е с т н о е движение обоих типов носителей называют *биполярной диффузией*. Если бы подвижности (а значит, и коэффициенты диффузии) у электронов и дырок были одинаковы, то они двигались бы в виде единого нейтрального потока, который можно было бы анализировать с помощью одного из уравнений (1-79). На самом деле подвижности носителей различны [см. (1-31)], поэтому у электронного потока

¹ Можно показать, что такое предположение соответствует условию пренебрежимости членами с $\partial E/\partial x$ в уравнениях непрерывности.

будет тенденция «обогнать» дырочный поток. В результате взаимного сдвига потоков образуются небольшой объемный заряд и соответствующее электрическое поле, которое тормозит поток электронов и ускоряет поток дырок.

В конце концов устанавливается стационарный режим, при котором избыточные электроны и дырки распределены в виде сдвинутых относительно друг друга «облачков». Эти «облачка» движутся синхронно, так что результирующий ток отсутствует. Описанные явления известны под названием *эффекта Дембера*, а электрическое поле и разность потенциалов, свойственные этому эффекту, называют *демберовским полем* и *демберовским напряжением*.

Значения демберовского поля и демберовского напряжения можно оценить из условия нулевого результирующего тока. Полагая $j = 0$ в формуле (1-71), подставляя значения токов из (1-72), (1-73) и используя соотношение (1-35), получаем:

$$E_{\text{демб}} = \frac{q}{\sigma} \left(D_p \frac{dp}{dx} - D_n \frac{dn}{dx} \right) \quad (1-100a)$$

или, если принять условие квазинейтральности ($dp/dx \approx dn/dx$),

$$E_{\text{демб}} = \frac{q}{\sigma} (D_p - D_n) \frac{dp}{dx}. \quad (1-100б)$$

Учитывая соотношение $-dU/dx = E_{\text{демб}}$, проинтегрируем (1-100a) в пределах от $x = 0$ до $x = \infty$ и положим $p(\infty) = p_0$ и $n(\infty) = n_0$. Тогда

$$U_{\text{демб}} = \frac{q}{\sigma} [D_p \Delta p(0) - D_n \Delta n(0)] \quad (1-101a)$$

или, если принять условие квазинейтральности ($\Delta p \approx \Delta n$),

$$U_{\text{демб}} = \frac{q}{\sigma} \Delta p(0) (D_p - D_n). \quad (1-101б)$$

Использование выражений (1-100) и (1-101) нуждается в предварительном определении избыточных концентраций и их произ-

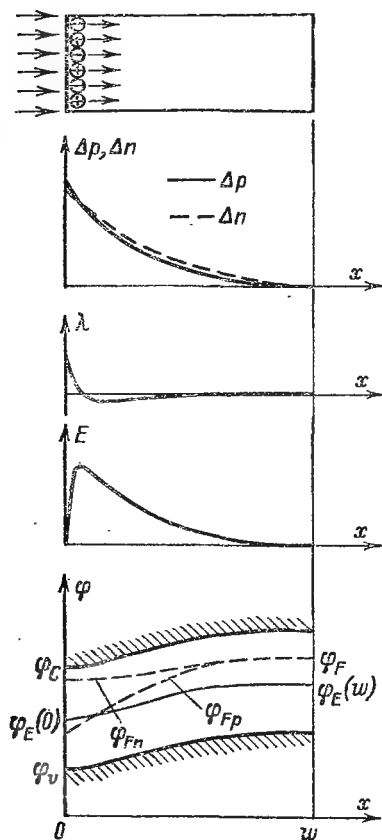


Рис. 1-31. Биполярная диффузия под действием света, генерирующего электроно-дырочные пары.

водных. Для этого нужно решить уравнения непрерывности (1-78), добавив в их правые части $\Delta g_p = \Delta g_n = \Delta g$. В условиях *полной* нейтральности члены с $\partial E/\partial x$ выпадают и оба уравнения становятся совершенно равноценными. Однако в условиях *квазинейтральности* априорное пренебрежение членами с $\partial E/\partial x$ рискованно, поскольку распределение носителей еще неизвестно и, следовательно, нельзя использовать критерий (1-99). Более корректно поступить иначе.

Умножим обе части (1-78а) на σ_n , а обе части (1-78б) — на σ_p . Тогда, если учесть (1-36б) и (1-36в), члены с $\partial E/\partial x$ будут одинаковыми с точностью до знака; складывая оба уравнения, можно от этих членов избавиться. В результате получается объединенное «электронно-дырочное» уравнение непрерывности, вид которого не зависит от наличия объемного заряда. В таком уравнении можно с полным основанием использовать условие квазинейтральности, приняв единое время жизни [см. (1-52)], а также положив одинаковыми избыточные концентрации и их производные. После этого, поделив обе части объединенного уравнения на $\sigma_n + \sigma_p = \sigma$, запишем его в следующей форме:

$$\frac{\partial \Delta p}{\partial t} = \Delta g - \frac{\Delta p}{\tau} + D_s \frac{\partial^2 \Delta p}{\partial x^2} + \mu_s E \frac{\partial \Delta p}{\partial x}. \quad (1-102)$$

Коэффициенты D_s и μ_s имеют следующие значения и названия:

$$D_s = \frac{D_n \sigma_p + D_p \sigma_n}{\sigma} \quad (1-103a)$$

— эффективный коэффициент диффузии или коэффициент биполярной диффузии;

$$\mu_s = \frac{\mu_n \sigma_p + \mu_p \sigma_n}{\sigma} \quad (1-103б)$$

— эффективная или биполярная дрейфовая подвижность¹.

В примесных полупроводниках, у которых $\sigma_n \gg \sigma_p$ или $\sigma_p \gg \sigma_n$, параметры биполярной диффузии D_s и μ_s практически совпадают с соответствующими параметрами для *основных* носителей. Это значит, что *процессы в примесных полупроводниках можно анализировать с помощью одного из уравнений (1-78), полагая $\partial E/\partial x = 0$.*

В собственных полупроводниках параметры D_s и μ_s существенно отличаются от соответствующих параметров для электронов и дырок. Это значит, что *при анализе процессов в собственном (или близком к собственному) полупроводнике нужно пользоваться уравнением (1-102).*

Заметим, что в собственном полупроводнике $\mu_s = 0$. Этот, на первый взгляд, парадоксальный факт объясняется тем, что под-

¹ Легко убедиться, что параметры D_s и μ_s не удовлетворяют соотношению (1-74). Поэтому иногда вводят еще понятие биполярной *диффузионной* подвижности, определяемой как D_s/φ_T .

близость μ_n к μ_p характеризует не движение носителей, а движение комплекса из электронного и дырочного «облачков» (см. выше). Поэтому при одинаковых концентрациях зарядов в «облачках» комплекс как целое нейтрален и не подвержен действию поля.

Возвращаясь к эффекту Дембера, можно констатировать, что определение поля $E_{\text{демб}}$ требует совместного решения уравнений (1-102) и (1-100б). При этом возникают чисто математические трудности (обусловленные нелинейностью дифференциального уравнения), на которых мы не будем останавливаться. Чтобы обойти эти трудности, стараются, если можно, рассматривать случай *малых возмущений* Δp , когда параметры σ , D , μ , τ сравнительно постоянны. Кроме того, пользуется распространением так называемое *диффузионное приближение*, при котором в уравнении (1-102) полагают $E = 0$, т. е. превращают его в диффузионное типа (1-79а), а затем, найдя решение, рассчитывают поле E по формуле (1-100б) и оценивают допустимость пренебрежения дрейфовой составляющей.

В заключение отметим, что биполярная диффузия, несмотря на отсутствие результирующего тока, является *неравновесным* процессом, так как имеются *потоки* носителей. Поэтому уровень Ферми не будет в данном случае постоянным, во всяком случае в той области, где потоки еще существенны и где повышены концентрации *обоих* типов носителей. В этих областях (см. рис. 1-31) уровень Ферми «расщепляется» на два квазиуровня — один для электронов, другой для дырок (см. § 1-6). Компоненты нулевого результирующего тока можно рассчитать по формулам (1-72) и (1-73), если рассмотренными выше методами определены распределения концентраций и поля.

Дрейф. В однородном полупроводнике внешнее приложенное напряжение порождает постоянное электрическое поле, не меняя при этом распределения концентраций носителей, т. е. не нарушая нейтральности (см. сноску на с. 74). Поэтому, суммируя дырочную и электронную составляющие (1-72), получаем плотность тока в виде обычной дифференциальной формы закона Ома:

$$j = (q\mu_p p + q\mu_n n) E = \sigma E. \quad (1-104)$$

Полный ток получается путем умножения обеих частей (1-104) на площадь поперечного сечения образца S . Поскольку напряженность однородного поля определяется соотношением $E = U/w$ (где U — приложенное напряжение и w — длина полупроводника), получается обычная форма записи закона Ома

$$U = IR, \quad (1-105)$$

где сопротивление образца

$$R = \frac{w}{\sigma S}. \quad (1-106)$$

Линейная зависимость $I(U)$ нарушается при достаточно больших напряжениях, когда поле в образце превышает критическое значение [см. (1-34)] и подвижности (а вместе с ними удельная проводимость) становятся функциями приложенного напряжения.

В неоднородных полупроводниках дело обстоит несколько сложнее. Будем считать площадь S неизменной. Тогда плотность тока во всех сечениях должна быть одинакова: $j = \text{const}$, а напряженность поля, обусловленного внешним напряжением, будет согласно (1-104) обратно пропорциональна удельной проводимости: $E(x) = j/\sigma(x)$. Иначе говоря, поле в неоднородном полупроводнике неоднородное. Если каждый элементарный участок неоднородного полупроводника рассматривать как однородный и считать для него действительным выражение (1-104), то для образца в целом оказывается действительным выражение (1-105), где сопротивление R является интегральной величиной:

$$R = \int_0^w \frac{dx}{\sigma(x) S}. \quad (1-107)$$

Эту формулу можно получить из соотношения $dU/dx = E$ (знак минус опущен, поскольку он не учитывается и в законе Ома):

$$U = \int_0^w E dx = \int_0^w \frac{j}{\sigma(x)} dx = j \int_0^w \frac{dx}{\sigma(x) S} = jR.$$

При необходимости в формуле (1-107) можно учесть переменное сечение пластинки $S(x)$.

Монополярная диффузия. Выше при анализе биполярной диффузии предполагалось, что оба типа носителей вводятся одновременно и в равных количествах. Поэтому «автоматически» обеспечивалась квазинейтральность, а результирующий ток отсутствовал.

Предположим теперь, что в приповерхностный слой полупроводника вводится (инжектируется) только один тип носителей — неосновных (рис. 1-32). Пусть для определенности осуществляется инжекция дырок в электронный полупроводник. Инжектированные дырки благодаря градиенту концентрации будут диффундировать в глубь кристалла, т. е. появится дырочный ток. Заряд дырок практически мгновенно (со временем диэлектрической релаксации, см. § 1-12) будет компенсирован таким же зарядом электронов, притягиваемых из глубоких слоев (необходимое количество дополнительных электронов поступает из внешней цепи, по которой протекает ток). В результате вблизи инжектирующей поверхности образуется квазинейтральное электронно-дырочное «облачко», почти такое же, как при биполярной диффузии (ср. рис. 1-31 и 1-32). Образуется также демберовское поле [см. (1-100)].

Процесс, при котором диффундируют носители одного типа, а носители другого типа лишь обеспечивают квазинейтральность, называют *монополярной диффузией*.

Несмотря на внешнее сходство в распределении носителей монополярная диффузия принципиально отличается от биполярной. Отличия состоят в следующем.

1. Имеется результирующий ток; поэтому, в частности, полупроводник должен быть элементом замкнутой цепи. Соответственно

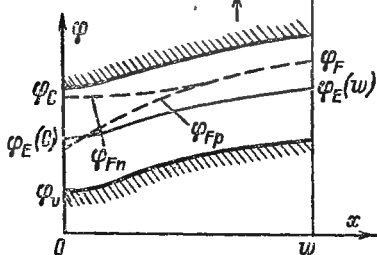
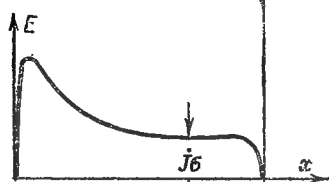
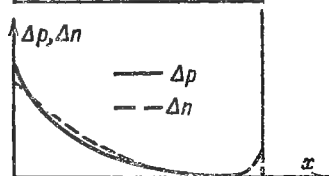
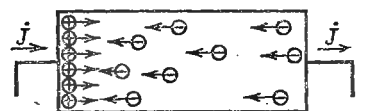


Рис. 1-32. Монополярная диффузия в результате инъекции дырок в электронный полупроводник.

помимо демберовского поля, сосредоточенного вблизи инжектирующей поверхности, во всей толще полупроводника действует «обычное» — *омическое* поле, обусловленное приложенным напряжением.

2. Потоки дырок и электронов направлены в разные стороны: дырки двигаются в глубь кристалла, а электроны — в сторону инжектирующей поверхности, в район электронно-дырочного «облачка», где происходит интенсивная рекомбинация и необходимо пополнение основных носителей. Результирующий ток является суммой дырочной и электронной составляющих.

3. В связи с постоянством результирующего тока дырочная и электронная составляющие меняются в разные стороны: убывание дырочного тока от поверхности в глубь кристалла сопровождается соответствующим ростом электронной составляющей. На самой инжектирующей поверхности электронный ток равен нулю, так как близко к нулю электрическое поле. Следовательно, в непосредственной близости от инжектирующей поверхности ток обусловлен только дырками. В глубине кристалла, где дырочный ток благодаря

рекомбинации делается равным нулю, обращается в нуль и диффузионная составляющая электронного тока. Следовательно, вдали от инжектирующей поверхности ток обусловлен только электронами и имеет чисто дрейфовый характер: электроны двигаются в омическом поле, созданном внешним напряжением.

Указанные отличия принципиальны с физической точки зрения. С математической же точки зрения распределения $\Delta p(x)$ и $\Delta n(x) \approx \Delta p(x)$ описываются либо уравнением биполярной диффузии (1-102) (если полупроводник близок к собственному), либо (в случае электронного полупроводника) уравнением (1-78a).

Что касается выражения для поля E , то оно в данном случае будет отличаться от (1-100б), поскольку результирующий ток не равен нулю. Подставляя (1-72) и (1-73) в (1-71), легко получить:

$$E = \frac{1}{\sigma} \left[j + q (D_p - D_n) \frac{dp}{dx} \right]. \quad (1-108)$$

Сравнивая (1-108) с (1-100б), замечаем, что поле при монополярной диффузии складывается из постоянной (омической) и демберовской составляющих. В глубоких, квазиравновесных слоях, где $dp/dx \approx 0$, остается омическое поле $E = j/\sigma$ в соответствии с выражением (1-104) для однородного полупроводника.

Строгое решение задачи о распределении носителей при монополярной диффузии затруднительно. Поэтому решим эту задачу для малых возмущений и в диффузионном приближении

Величину возмущения принято характеризовать *уровнем инжекции*

$$\delta = \frac{\Delta p}{n_0} \approx \frac{\Delta n}{n_0},$$

т. е. отношением концентрации избыточных носителей к равновесной концентрации основных носителей. Малому возмущению соответствует неравенство $\delta \ll 1$, т. е.

$$\Delta p \approx \Delta n \ll n_0, \quad (1-109)$$

которое называют условием *низкого уровня инжекции*. Это условие дает право пренебречь эффектом модуляции проводимости на возмущенном участке и принять в формуле (1-103а) соотношение $\sigma \approx \sigma_n \gg \sigma_p$, характерное для равновесного электронного полупроводника. Следствием этого можно считать равенство $D_s = D_p$.

Диффузионное приближение характеризуется тем, что в уравнении (1-102) пренебрегают членом с E . С учетом этого пренебрежения, а также условия $D_s = D_p$ уравнение (1-102) переходит в диффузионное уравнение (1-79а).

Поделив обе части (1-79а) на D_p , запишем нестационарный и стационарный варианты диффузионного уравнения в следующем виде:

$$\frac{\partial^2 \Delta p}{\partial x^2} - \frac{\Delta p}{L_p^2} = \frac{\tau_p}{L_p^2} \frac{\partial \Delta p}{\partial t}; \quad (1-110a)$$

$$\frac{d^2 \Delta p}{dx^2} - \frac{\Delta p}{L_p^2} = 0. \quad (1-110б)$$

Величина L (индекс для общности опущен) определяется выражением

$$L = \sqrt{D\tau}; \quad (1-111)$$

она имеет размерность длины и носит название *средней диффузионной длины* или *средней длины диффузии*. Физический смысл величины L будет ясен из дальнейшего.

Остановимся сначала на стационарном уравнении (1-110б). Как известно, его решение является суммой двух экспонент:

$$\Delta p(x) = A_1 e^{-x/L_p} + A_2 e^{x/L_p}, \quad (1-112)$$

где коэффициенты A_1 и A_2 определяются из граничных условий (обычно при $x = 0$ и $x = w$). Если $w = \infty$ и $\Delta p(\infty) = 0$, то $A_2 = 0$ и $A_1 = \Delta p(0)$.

Тогда

$$\Delta p(x) = \Delta p(0) e^{-x/L_p}, \quad (1-113)$$

т. е. избыточные концентрации спадают по экспоненте. Это один из характерных случаев в теории полупроводниковых приборов.

Из выражения (1-113) следует, что *диффузионная длина есть то расстояние, на котором концентрация диффундирующих носителей (при их экспоненциальном распределении) уменьшается в e раз*. На участке длиной $(3 \div 4) L$ концентрация уменьшается в 20—50 раз, т. е. становится пренебрежимо малой по сравнению с исходной.

Дифференцируя (1-113), получаем градиент концентрации:

$$\frac{dp}{dx} = -\frac{\Delta p(x)}{L_p}. \quad (1-114a)$$

На инжектирующей плоскости, т. е. в точке $x = 0$,

$$\frac{dp}{dx}(0) = -\frac{\Delta p(0)}{L_p}. \quad (1-114б)$$

Как видим, при экспоненциальном распределении носителей градиент их концентрации (в том числе начальный) пропорционален избыточной концентрации в каждой данной точке.

В процессе диффузии избыточные носители (в нашем случае — дырки) приобретают определенную скорость, которую можно оценить из соотношения (см. сноску * на с. 42):

$$v = j/\lambda,$$

где $\lambda = q \Delta p$ — плотность заряда избыточных дырок.

Используя выражения (1-73а), (1-114а) и опуская для общности индекс p , получаем *диффузионную скорость* в виде

$$v_{\text{диф}} = \frac{D}{L} = \frac{L}{\tau}. \quad (1-115)$$

Следовательно, при экспоненциальном распределении носителей их диффузионная скорость является постоянной величиной. Исходя из выражения (1-115), можно дать еще одно определение величины L : *диффузионная длина есть то среднее расстояние, которое носители в процессе диффузии проходят за время жизни.*

Если подставить dp/dx из (1-114а) в (1-73а), получим выражение для диффузионной составляющей дырочного тока. Запишем ее в следующем виде:

$$(j_p)_{\text{диф}}(x) = (j_p)_{\text{диф}}(0) e^{-x/L_p}, \quad (1-116a)$$

где граничное значение в точке $x = 0$

$$(j_p)_{\text{диф}}(0) = -qD_p \frac{\Delta p(0)}{L_p}. \quad (1-116b)$$

Заметим, что это граничное значение совпадает с полным током:

$$(j_p)_{\text{диф}}(0) = j, \quad (1-117)$$

поскольку при $x = 0$ электронный ток отсутствует (см. начало данного раздела, п. 3), а дрейфовой составляющей дырочного тока мы пренебрегли с самого начала, приняв «диффузионное приближение».

Подставляя dp/dx из (1-114а) в (1-108) и учитывая выражения (1-113), (1-116b) и (1-117), получаем напряженность того поля, которым пренебрегли при решении уравнения непрерывности. Если использовать обозначение (1-31), то

$$E(x) = \frac{j}{\sigma} [1 + (b-1)e^{-x/L_p}]. \quad (1-118)$$

Граничные значения поля:

$$E(0) = b(j/\sigma); \quad E(\infty) = j/\sigma.$$

Напряженность демберовской составляющей поля определяется вторым слагаемым в формуле (1-118):

$$E_{\text{демб}} = \frac{j}{\sigma} (b-1)e^{-x/L_p}, \quad (1-119)$$

граничные значения этой составляющей имеют вид:

$$E_{\text{демб}}(0) = (b-1)(j/\sigma); \quad E_{\text{демб}}(\infty) = 0.$$

Распределение потенциала при монополярной диффузии несколько отличается от распределения при биполярной диффузии из-за различия полей. В частности, на расстоянии $(2 \div 3)L$ от плоскости инжекции, т. е. в области квазиравновесия, потенциал меняется почти линейно (ср. с рис. 1-17, б). Квазиуровни Ферми, определяемые избыточными концентрациями, меняются практически так же, как при биполярной диффузии.

Нестационарное уравнение (1-110а) может быть решено разными способами. В инженерной практике наиболее распространен операторный метод (см. [22]). Как известно, при этом методе функция времени, в нашем случае $\Delta p(x, t)$, заменяется ее операторным изображением $\Delta p(x, s)$, а производная по времени — произведением¹

$$s[\Delta p(x, s) - \Delta p(x)_{t=0}]. \quad (1-120)$$

¹ Здесь и ниже оператор Лапласа обозначается через s , а не через p (как обычно) во избежание путаницы с обозначением концентрации дырок.

После этого при решении диффузионного уравнения оператор рассматривается как алгебраический множитель, а найденное решение типа (1-112) оказывается *изображением* искомой функции $\Delta p(x, t)$. Эта функция (*оригинал изображения*) в большинстве случаев может быть найдена по таблицам соответствия.

Полагая начальное состояние равновесным, т. е. $\Delta p(x)_{t=0} = 0$, нетрудно привести уравнение (1-110a) к виду

$$\frac{d^2 \Delta p}{dx^2} - \frac{1 + s\tau_p}{L_p^2} \Delta p = 0. \quad (1-121)$$

Как видим, отличие этой формулы от (1-110б) состоит только в том, что диффузионная длина должна рассматриваться как операторная величина

$$L(s) = \frac{L}{\sqrt{1 + s\tau}}. \quad (1-122)$$

Оригинал изображения $L(s)$ имеет вид (рис. 1-33):

$$L(t) = L \operatorname{erf} \left(\frac{t}{\tau} \right), \quad (1-123)$$

где $\operatorname{erf}(t/\tau)$ — функция ошибок¹.

Необходимо заметить, что определение диффузионной длины (см. выше) связано с экспоненциальным распределением носителей, тогда как во время переходного процесса (особенно в его начальной части) распределение неэкспоненциальное. Поэтому выражение $L(t)$ нельзя механически подставлять вместо L в «готовые» формулы (1-112) и (1-113). Тем не менее, рассматривая диффузионную длину как функцию времени, можно делать полезные качественные выводы о характере переходных процессов.

Учитывая структурное сходство уравнений (1-121) и (1-110б), приходим к выводу, что решения (1-112) и (1-113) остаются в силе, если заменить L_p на $L_p(s)$. Граничные условия для определения коэффициентов A_1 и A_2 используются так же, как в стационарном случае. В результате получается изображение $\Delta p(x, s)$, по которому с помощью таблиц можно найти оригинал $\Delta p(x, t)$. Такой оригинал будет рассмотрен в § 2-9 в связи с переходным процессом в диоде.

¹ Функция ошибок $\operatorname{erf}(z)$ (ее же называют *интегралом ошибок* или *функцией Крампса*) определяется следующим образом [23]:

$$\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-\xi^2} d\xi.$$

Согласно рис. 1-33 функция $\operatorname{erf}(z)$ в общем напоминает экспоненциальную функцию $1 - e^{-z}$; в частности, $\operatorname{erf}(0) = 0$ и $\operatorname{erf}(\infty) = 1$. Производная от $\operatorname{erf}(z)$ имеет вид $(2/\sqrt{\pi}) e^{-z^2}$. *Дополнительной функцией ошибок* называют функцию $\operatorname{cerf}(z) = 1 - \operatorname{erf}(z)$, которая напоминает экспоненту e^{-z} . Производная функции $\operatorname{cerf}(z)$ отличается от производной функции $\operatorname{erf}(z)$ только знаком.

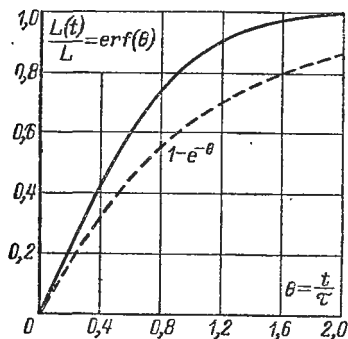


Рис. 1-33. Зависимость диффузионной длины от времени [функция ошибок $\operatorname{erf}(t/\tau)$].

Весь материал данного раздела непосредственно относится к случаю инжекции, когда $\Delta p > 0$. Однако с математической точки зрения анализ остается в силе и для случая экстракции, когда $\Delta p < 0$. Нужно лишь иметь в виду следующие физические особенности, свойственные режиму экстракции.

1. Концентрации носителей в приповерхностном слое оказываются меньше равновесных; соответственно меняются знак объемного заряда, направление полей, а также направления потоков носителей.

2. Поскольку полная концентрация дырок p не может быть отрицательной, то модуль отрицательной избыточной концентрации всегда меньше равновесной концентрации: $|\Delta p| < p_0$. Отсюда следует, что в режиме экстракции возмущения всегда малы; соответственно малы плотности токов (токи).

Комбинированное движение. В последних двух разделах дрейф и диффузия носителей рассматривались отдельно: в первом случае считалось, что нет инжекции, во втором — поле считалось пренебрежимо малым.

Проанализируем диффузию в условиях существенного электрического поля; для простоты положим напряженность поля постоянной¹. Такой случай имеет, например, место при инжекции в неоднородный полупроводник с экспоненциальным распределением примеси (1-93). Будем, как и раньше, рассматривать инжекцию дырок в электронный полупроводник. Тогда в уравнении (1-102) следует положить $D_s = D_p$, $\mu_s = -\mu_p$, а также $\Delta g = 0$. Ограничиваясь стационарным режимом ($dp/dt = 0$), поделим обе части уравнения на D_p и запишем его в следующей форме:

$$\frac{d^2 \Delta p}{dx^2} - \frac{\mu_p E}{D_p} \frac{d \Delta p}{dx} - \frac{\Delta p}{L_p^2} = 0.$$

Величина $\mu_p E/D_p$ имеет размерность 1/см, поэтому ее можно представить как $2\eta/L_p$, где безразмерный коэффициент η с учетом (1-115) характеризует соотношение дрейфовой и диффузионной скоростей, а с учетом (1-74) — относительную роль напряженности поля²:

$$\eta = \frac{\mu E}{2D/L} = \frac{E}{2\varphi_T/L}. \quad (1-124)$$

Обычно величину η называют *коэффициентом поля*. С использованием коэффициента поля уравнение непрерывности запишется в виде

$$\frac{d^2 \Delta p}{dx^2} - \frac{2\eta}{L_d} \frac{d \Delta p}{dx} - \frac{\Delta p}{L_p^2} = 0. \quad (1-125)$$

¹ В противном случае, как уже отмечалось, задача делается нелинейной.

² Величина φ_T/L , будучи умножена на подвижность μ , дает диффузионную скорость $v_{\text{диф}}$ (1-115). Поэтому φ_T/L можно назвать *диффузионной напряженностью поля*.

Известно, что решение такого уравнения является суммой двух экспоненциальных членов, как и (1-112), но показатели экспонент имеют значения $x\xi_1$ и $x\xi_2$, где ξ_1 и ξ_2 — корни характеристического уравнения (в данном случае это полное квадратное уравнение). Учитывая, что корни имеют размерность 1/см, их можно записать следующим образом:

$$\xi_{1,2} = \frac{1}{\mathcal{L}_{1,2}} = \frac{1}{L} (\eta \mp \sqrt{\eta^2 + 1}), \quad (1-126)$$

где величина $\mathcal{L}$ является аналогом диффузионной длины и носит название *глубины затягивания*.

Один из корней (1-126) положителен: $\xi_2 > 0$. Поэтому, если полупроводник имеет бесконечную длину ($w = \infty$), то коэффициент A_2 в выражении (1-112) должен быть равен нулю и, следовательно, положительный корень не представляет интереса. Тогда решение уравнения непрерывности запишется в виде

$$\Delta p(x) = \Delta p(0) e^{-x/L}, \quad (1-127)$$

где глубина затягивания

$$\mathcal{L} = \frac{L}{\sqrt{\eta^2 + 1} - \eta}. \quad (1-128)$$

Как видим, глубина затягивания больше диффузионной длины, если $\eta > 0$, т. е. если $E > 0$. Это понятно, поскольку положительное поле способствует движению дырок в направлении диффузии. Если $E < 0$ и соответственно $\eta < 0$, то поле будет тормозить дырки и глубина затягивания будет меньше диффузионной длины. В случае экстракции дырок, когда направления потоков меняются, положительное поле тормозит дырки, т. е. их распределение растягивается ($\mathcal{L} > L$), а отрицательное поле ускоряет дырки, т. е. «стягивает» их к экстрагирующей поверхности ($\mathcal{L} < L$). Рисунок 1-34 иллюстрирует эти явления, два из которых получили специальные названия: *экслюзия* и *аккумуляция* носителей.

Экслюзией называют обеднение носителями того слоя, который при чистой диффузии был бы обогащен благодаря инъекции (кривая 3). Аккумуляцией называют обогащение носителями того слоя, который при чистой диффузии был бы обеднен благодаря экстракции (кривая 6).

Из выражения (1-128) следует, что если $\eta \ll 1$ (практически, если $\eta < 0,2 \div 0,3$), то влияние поля незначительно, т. е. $\mathcal{L} \approx L$.

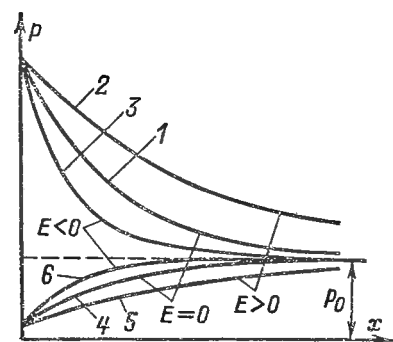


Рис. 1-34. Распределение дырок при инъекции (кривые 1—3) и экстракции (кривые 4—6) в условиях отсутствия и наличия электрических полей разной полярности.

Наоборот, если $n \gg 1$ (практически, если $\eta > 2$), то движение носителей будет почти чисто дрейфовым. В этом случае, используя преобразование

$$\sqrt{\eta^2 + 1} = \eta \sqrt{1 + \eta^{-2}} \approx \eta (1 + \eta^{-2}/2),$$

получаем с учетом (1-111):

$$\mathcal{L} \approx 2\eta L = \mu E \frac{L^2}{D} = v_{\text{др}} \tau,$$

т. е. глубина затягивания оказывается равной расстоянию, которое неосновные носители проходят со скоростью дрейфа за время жизни.

Конечно, все сделанные выводы действительны и для комбинированного движения электронов в дырочном полупроводнике с учетом того, что ускоряющим для электронов является отрицательное поле.

Глава вторая

ПОЛУПРОВОДНИКОВЫЕ ДИОДЫ

2-1. ВВЕДЕНИЕ

Полупроводниковый диод представляет собой комбинацию двух полупроводниковых слоев с различными типами проводимости (рис. 2-1, а).

Такая комбинация обладает способностью гораздо лучше пропускать ток в одном направлении (от слоя p к слою n) и гораздо хуже в другом. Полярность напряжения, соответствующая большим токам, называется прямой, а соответствующая меньшим токам — обратной. Обычно пользуются терминами *прямое* и *обратное напряжение*, *прямой* и *обратный ток*. На рис. 2-1, б показаны символическое изображение диода, направление прямого тока и полярность прямого напряжения.

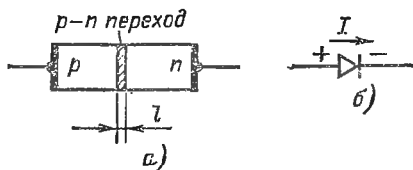


Рис. 2-1. Плоскостной диод.

а — упрощенная структура; б — условное обозначение.

Вентильное свойство диода является следствием ярко выраженной внутренней неоднородности структуры. Ступенчатая неоднородность даже в полупроводниках с одним типом проводимости сопровождается нарушением закона Ома в связи с образованием объемных зарядов и потенциального барьера (см. рис. 1-30, б). В данном случае, когда слои разнотипные, нелинейность, естественно, оказывается еще сильнее.

Поверхность, по которой контактируют слои p и n , называется *металлургической границей*, а прилегающая к ней область объемных (или пространственных) зарядов — *электронно-дырочным пере-*

ходом или *p-n* переходом. Два других внешних контакта в диоде не выпрямляющие: поэтому их называют *омическими контактами*.

Плоскостными диодами и соответственно плоскостными переходами называют такие диоды и переходы, у которых граница между слоями плоская, а площади обоих слоев одинаковы. Эти условия не всегда соблюдаются на практике, однако они облегчают анализ и в то же время позволяют получить правильное представление о процессах в реальном диоде и его характеристиках.

Диоды — самостоятельный, весьма обширный класс полупроводниковых приборов. В то же время диод как простейший прибор с одним *p-n* переходом является основой многопереходных приборов — транзисторов. Мы изучим диоды именно с этой точки зрения, уделив главное внимание тем процессам и свойствам, которые характерны как для диодов, так и для транзисторов.

В общих чертах процессы в полупроводниковом диоде можно охарактеризовать следующим образом. В отсутствие внешнего напряжения в *p-n* переходе имеет место больцмановское равновесие: диффузионные и дрейфовые потоки носителей уравновешены и результирующего тока нет.

Полярность прямого напряжения способствует «выталкиванию» дырок из *p*-слоя в *n*-слой и электронов в обратном направлении, т. е. инжекции неосновных носителей в оба слоя. Инжектированные носители диффундируют в глубь слоев, и эта монополярная диффузия сопровождается протеканием достаточно большого прямого тока (см. § 1-13).

Полярность обратного напряжения способствует «выталкиванию» дырок из *n*-слоя и электронов из *p*-слоя в область перехода, т. е. экстракции неосновных носителей. При этом протекает небольшой обратный ток.

Согласно (1-116) диффузионный ток определяется величиной граничной избыточной концентрации неосновных носителей. Поскольку при экстракции избыточная концентрация не может (по модулю) превысить весьма малую равновесную концентрацию (см. с. 82, п. 2), а при инжекции такого ограничения нет, то прямой ток оказывается намного больше обратного. Это и является основой вентильных свойств диода.

2-2. ЭЛЕКТРОННО-ДЫРОЧНЫЙ ПЕРЕХОД

Плоскостной диод состоит из электронно-дырочного перехода, двух нейтральных (или квазинейтральных) слоев и омических контактов. Поскольку процессы в нейтральных полупроводниках были детально изучены в гл. 1, следует прежде всего рассмотреть процессы в *p-n* переходах. Свойства омических контактов будут описаны позднее.

Классификация *p-n* переходов. Прежде всего заметим, что *p-n* переход нельзя осуществить путем простого соприкосновения двух разнородных полупроводниковых пластинок, так как при

этом неизбежен промежуточный слой (хотя бы и очень тонкий) слой воздуха или поверхностных пленок. Настоящий переход получается в единой пластинке полупроводника, в которой тем или иным способом (см. § 4-13) получена достаточно резкая граница между слоями p и n .

Резкость границы играет существенную роль для формирования перехода, так как чересчур плавный переход, как показывает теория, не обладает теми вентиляльными свойствами, которые лежат в основе работы полупроводниковых диодов и транзисторов [24]. Понятие резкости формулируется следующим образом: граница между слоями является резкой, если градиент концентрации примеси (считающийся постоянным в пределах перехода) удовлетворяет неравенству

$$\left| \frac{dN}{dx} \right| l_{Di} \gg n_i, \quad (2-1)$$

где N — эффективная концентрация примеси (см. стр. 21), а l_{Di} — дебаевская длина в собственном полупроводнике [см. (1-87)]. Для кремния и германия необходимо соответственно: $dN/dx \gg 10^{13} \text{ см}^{-4}$ и $dN/dx \gg 4 \cdot 10^{17} \text{ см}^{-4}$.

Критерий (2-1) говорит о том, что концентрация примесей в переходе должна существенно изменяться на отрезке, меньшем l_{Di} . Такое требование имеет определенный физический смысл. Пусть $N = N_a - N_d$ есть эффективная концентрация акцепторов, которая линейно спадает в направлении слоя n (рис. 2-2). В точке $x = 0$, где $N = 0$, имеем компенсированный, т. е. практически собственный полупроводник (см. рис. 1-13, в). Далее, при $x > 0$, нарастает эффективная концентрация доноров с прежним (по модулю) градиентом. Если градиент dN/dx не удовлетворяет условию (2-1), т. е. очень мал, то существенные (сравнимые с n_i) эффективные концентрации примесей получаются вдали от точки $x = 0$, на расстояниях, значительно больших глубины экранирования l_{Di} . Тогда поля объемных зарядов, обусловленных ионизированными примесными атомами, тоже будут расположены вдали от металлургической границы и не будут «соприкасаться» в точке $x = 0$. Соответственно в окрестности металлургической границы не сможет образоваться двойной электрический слой, свойственный электронно-дырочному переходу.

Переходы, в которых имеется скачкообразное изменение концентрации на границе слоев ($dN/dx = \infty$), будем называть *ступенчатыми*. Они представляют собой предельный случай более общего класса *плавных* переходов, в которых градиент концентрации примесей конечен, но удовлетворяет неравенству (2-1). На практике ступенчатые переходы являются, конечно, известным приближением. Однако они хорошо отражают свойства многих реальных

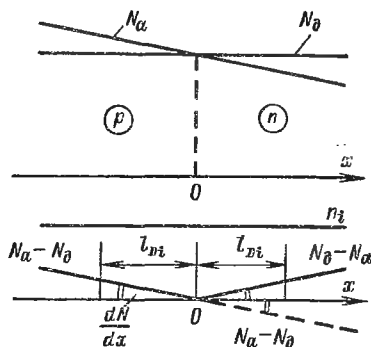


Рис. 2-2. Распределение полных и эффективных концентраций примеси вблизи металлургической границы плавного перехода.

p-n структур и, кроме того, оказываются проще для анализа. Поэтому ниже им будет уделено главное внимание. Контакты, в которых условие (2-1) не соблюдается, не называют переходами, а относят к неоднородным полупроводникам.

По соотношению концентраций основных носителей в слоях *p* и *n* переходы делятся на *симметричные* и *несимметричные*. В симметричных переходах имеет место соотношение

$$p_p \approx n_n,$$

т. е. концентрации основных носителей в обоих слоях почти одинаковы. Такие переходы трудно реализовать практически, и они не являются типичными. Гораздо большее распространение имеют несимметричные переходы, в которых выполняется неравенство

$$p_p > n_n \text{ или } n_n > p_p$$

и концентрации различаются в несколько раз и более. Именно такие переходы будут анализироваться в дальнейшем, причем для определенности будет считаться, что слой *p* более низкоомный, чем слой *n*, т. е. $p_p > n_n$. Полученные выводы легко использовать при обратном соотношении концентраций.

В случае резкой асимметрии, когда концентрации основных носителей различаются более чем на порядок, переходы называют *односторонними* и обычно обозначают символами p^+-n (или n^+-p). Иногда, чтобы отличить «просто» несимметричные переходы от односторонних, используют для первых обозначения $p^{++}-n$ (или $n^{++}-p$), а для вторых $p^{+-}n$ (или $n^{+-}p$).

В дальнейшем несимметричность (и даже односторонность) переходов будет подразумеваться без использования указанных индексов, за исключением специальных случаев, когда отсутствие индексов может вызвать недоразумения.

Структура *p-n* перехода. Концентрации примесей и свободных носителей в каждом из слоев диода показаны на рис. 2-3, а, причем для наглядности разница в концентрациях p_p и n_n принята гораздо меньшей, чем это имеется в действительности.

Поскольку концентрация дырок в слое *p* значительно больше, чем в слое *n*, часть дырок диффундирует из слоя *p* в слой *n*. При этом в слое *n* вблизи границы окажутся избыточные дырки, которые будут рекомбинировать с электронами до тех пор, пока не будет выполнено условие равновесия (1-16). Соответственно в этой области уменьшится концентрация свободных электронов и «обнажатся» некомпенсированные положительные заряды донорных атомов. Слева от границы «обнажатся» некомпенсированные отрицательные заряды акцепторных атомов, поскольку часть дырок перешла отсюда в слой *n* (рис. 2-3, б). Аналогичные рассуждения действительны для электронов слоя *n*, которые частично диффундируют в слой *p*. Однако в несимметричном переходе, в котором $n_n \ll p_p$, диффузия электронов в слой *p* малосущественна, поскольку разность концентраций $n_n - n_p$ несравненно меньше разницы $p_p - p_n$, а именно этими разно-

стями согласно (1-114) определяются градиенты концентраций и диффузионные токи.

Область образовавшихся пространственных зарядов и есть область p - n перехода. Часто эту область называют *обедненным* или *истощенным* слоем, имея в виду резко пониженную концентрацию подвижных носителей в обеих ее частях. Однако, строго говоря, переход и обедненный слой — не одно и то же: область перехода несколько шире потому что объемные заряды и связанное с ними поле зарождаются уже при очень небольшом

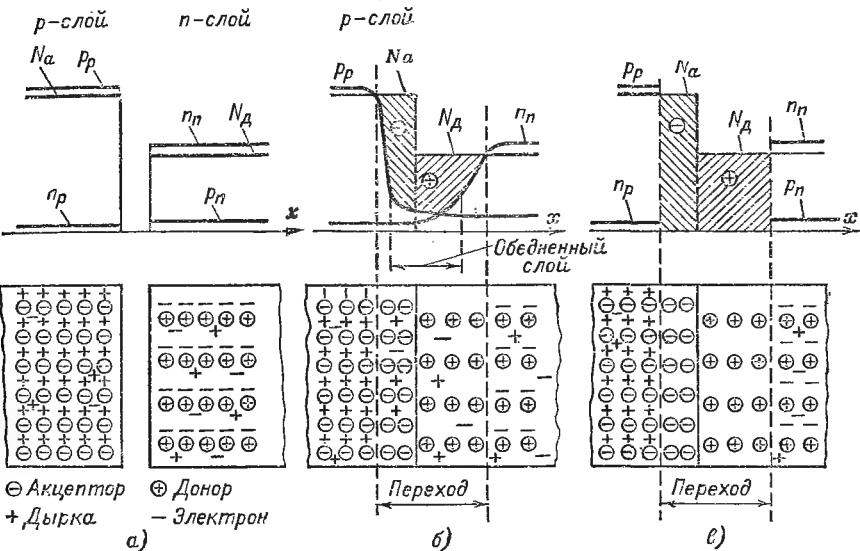


Рис. 2-3. Структура p - n перехода.

a — начальное состояние слоев; *б* — пространственные заряды в реальном переходе; *в* — пространственные заряды в идеализированном переходе.

(несколько процентов) уменьшении концентрации носителей по сравнению с равновесной, тогда как понятию обедненного слоя соответствует спад концентрации носителей по крайней мере на порядок (рис. 2-3, б). Промежуточные участки между «границами» обедненного слоя и перехода являются участками экранирования p - и n -слоев диода от поля создаваемого зарядами обедненного слоя. При перепаде концентраций носителей на три порядка и более протяженность участков экранирования обычно не превышает 0,1 мкм, тогда как ширина собственно обедненного слоя, как увидим ниже, в несколько раз больше. Поэтому есть основания идеализировать переход так, как показано на рис. 2-3, в, т. е. полностью пренебречь наличием свободных носителей в переходе и считать его границы совпадающими с границами обедненного слоя. Такая идеализация существенно упрощает решение многих задач, за исключением, конечно, тех, которые непосредственно связаны с анализом потоков носителей.

Переход в целом, нейтрален, т. е. отрицательный заряд в левой части и положительный заряд в правой части одинаковы. При этом условии различие в концентрациях акцепторной и донорной примесей неизбежно связано с различием в протяженности обоих зарядов: в слое с меньшей концентрацией примеси (в нашем случае в n -слое) область объемного заряда должна быть шире. Иначе говоря, *несимметричный переход сосредоточен в высокоомном слое*. Учет зарядов дырок и электронов в переходе [25] не меняет этого важного вывода.

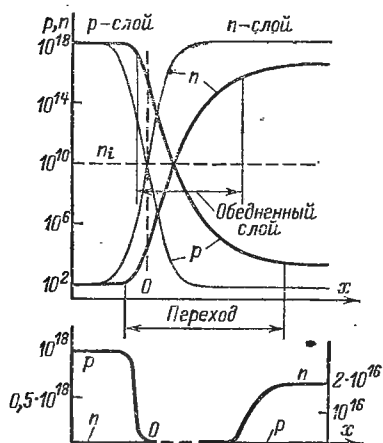


Рис. 2-4. Распределение носителей в несимметричном переходе (полулогарифмический и линейный масштаб). Тонкими линиями показано распределение в симметричном переходе.

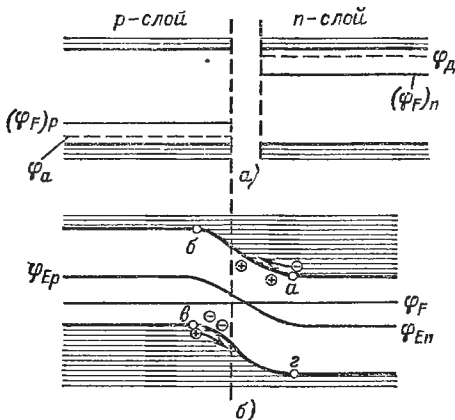


Рис. 2-5. Зонная диаграмма слоев (а) и p - n перехода в равновесном состоянии (б).

На рис. 2-4 показано распределение носителей в полулогарифмическом масштабе, более удобном для количественных оценок и сравнений [26]. Особое внимание следует обратить на тот факт, что внутри p - n перехода имеется участок с собственной (т. е. сильно пониженной по отношению к основным слоям) проводимостью. Тем самым область перехода является наиболее высокоомной частью диодной структуры. Участок с собственной проводимостью не совпадает с металлургической границей, а сдвинут в сторону того слоя, где сосредоточен переход.

Пространственные заряды в переходе образуют электрическое поле, которое направлено так, что оно ограничивает диффузию носителей. В равновесном состоянии диффузионные потоки носителей, обусловленные градиентами концентрации, в любой точке равны, но направлены навстречу дрейфовым потокам тех же носителей, обусловленным градиентом потенциала (больцмановское равновесие).

Рассмотрим теперь p - n переход с точки зрения зонной теории. В отсутствие контакта совокупность p - и n -слоев характеризуется диаграммой на рис. 2-5, *а*. При наличии контакта уровень Ферми должен быть единым, а это приводит к неизбежному искривлению зон, различию электростатических потенциалов и образованию потенциального барьера (рис. 2-5, *б*).

Если воспользоваться образной интерпретацией движения носителей в зонах (см. сноску на с. 37), то равновесное состояние перехода можно охарактеризовать следующим образом. Основная масса дырок p -слоя диффундирует слева направо в область перехода, но не может преодолеть потенциальный барьер и, проникнув в переход на некоторую глубину, «отражается» и возвращается в p -слой (рис. 2-6). Дырки же n -слоя независимо от энергии беспрепятственно «всплывают» в p -слой и образуют поток справа налево. Этот поток уравнивается встречным потоком достаточно энергичных дырок p -слоя, способных преодолеть барьер. Аналогичная ситуация имеет место по отношению к электронам: электроны p -слоя свободно «скатываются» в n -слой. Этот поток уравнивается потоком наиболее энергичных электронов n -слоя. Основная же масса электронов этого слоя, «пытающаяся» диффундировать в p -слой, отражается потенциальным барьером (рис. 2-6). Глубина проникания отражаемых носителей в переход тем больше, чем выше их энергия.

В области перехода на рис. 2-5, *б* показаны ионизированные атомы акцепторов (слева) и доноров (справа). Как известно, уровни этих ионов расположены вдоль всего соответствующего слоя, но мы показываем их только в пределах перехода, чтобы подчеркнуть некомпенсированность заряда ионов на этих участках. В самом деле, расстояние между дном зоны проводимости и уровнем Ферми увеличивается от точки *а* влево, а значит, на участке *а—б* быстро убывает вероятность заполнения зоны проводимости электронами. Поэтому справа от точки *а* электроны компенсируют положительный заряд донорных ионов и слой n нейтрален, а слева от точки *а* концентрация электронов резко падает и такой компенсации нет. Аналогично обстоит дело на участке справа от точки *в* по отношению к акцепторным ионам. Очевидно, что ионы, показанные на рис. 2-5, *б*, соответствуют ионам, образующим пространственный заряд на рис. 2-3.

Анализ перехода в равновесном состоянии. Анализ p - n переходов в общем весьма сложен. Наиболее просто анализируется ступенчатый переход, рассмотренный в предыдущем разделе. Но даже в этом случае необходимы упрощающие предположения. В самом деле, если решать задачу строго, т. е. исходить из структуры пере-

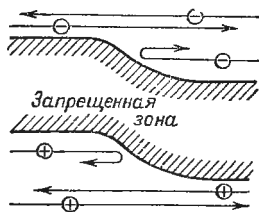


Рис. 2-6. Зонная схема движения носителей в переходе.

хода, показанной на рис. 2-3, б, то нужно учитывать распределение подвижных носителей в переходе. Тогда функция $\varphi_E(x)$, по которой можно найти все остальные интересующие величины, определяется уравнением¹

$$\frac{\partial^2 \Phi}{\partial x^2} = \frac{1}{l_{Di}^2} \left(\text{sh } \Phi + \frac{1}{2} \frac{N}{n_i} \right), \quad (2-2)$$

где $N = -N_a$ в p -слое и $N = N_d$ в n -слое.

Такое уравнение не имеет аналитического решения. Поэтому задачу упрощают, предполагая, что переход имеет структуру, показанную на рис. 2-3, в, где концентрации подвижных носителей внутри перехода равны нулю. Тогда в правой части уравнения (2-2) исчезнет первое слагаемое, обусловленное свободными носителями, и решение становится элементарным [см. (2-6) и (2-7)].

Если пока отвлечься от формы потенциального барьера, т. е. от функции $\varphi_E(x)$, то высоту равновесного потенциального барьера можно получить непосредственно из рис. 2-5, б:

$$\Delta \varphi_0 = \varphi_{Ep} - \varphi_{En}, \quad (2-3)$$

где величины в правой части — электростатические потенциалы в глубине p - и n -областей. Воспользуемся формулой (1-18а) и запишем эти потенциалы через концентрации свободных электронов в p - и n -слоях:

$$\varphi_{Ep} = -\varphi_T \ln \frac{n_{p0}}{n_i} + \varphi_F;$$

$$\varphi_{En} = -\varphi_T \ln \frac{n_{n0}}{n_i} + \varphi_F,$$

где индекс 0 соответствует равновесному состоянию.

Подставляя эти значения в (2-3), приходим к выражению

$$\Delta \varphi_0 = \varphi_T \ln \frac{n_{n0}}{n_{p0}}. \quad (2-4а)$$

Если воспользоваться формулой (1-18б) или в выражении (2-4а) заменить концентрации электронов концентрациями дырок с помощью соотношения (1-16), то высота потенциального барьера запишется следующим образом:

$$\Delta \varphi_0 = \varphi_T \ln \frac{p_{p0}}{p_{n0}}. \quad (2-4б)$$

Величину $\Delta \varphi_0$ иногда называют *диффузионным потенциалом*, поскольку эта разность потенциалов, во-первых, образуется в результате диффузии носителей через переход и, во-вторых, противодействует диффузионным потокам носителей. Еще

¹ Уравнение (2-2) является обобщением уравнения (1-86) на случай примесного полупроводника, когда плотность объемного заряда определяется выражением (1-81а).

одно название для величин $\Delta\Phi_0$ — *контактная разность потенциалов* — ставит ее в один ряд с аналогичной величиной, хорошо известной из теории металлических контактов; при этом, однако, не отражается специфика p - n перехода, состоящая в том, что он не может быть получен путем спайки или сварки соответствующих компонентов.

Выражая в любой из последних формул концентрацию неосновных носителей (n_p или p_n) через концентрацию основных носителей в том же слое с помощью соотношения (1-16) и используя выражения (1-36), легко выразить высоту равновесного барьера через удельные сопротивления слоев:

$$\Delta\Phi_0 = \varphi_T \ln \frac{\rho_i^2 (b+1)^2}{\rho_p \rho_n b}, \quad (2-4в)$$

где $b = \mu_n/\mu_p$ (1-31).

В качестве примера возьмем значения ρ_i и μ для кремния (табл. 1-1) и примем $\rho_p = 0,001 \text{ Ом} \cdot \text{см}$ и $\rho_n = 1 \text{ Ом} \cdot \text{см}$. Тогда при комнатной температуре, когда $\varphi_T = 0,025 \text{ В}$, получим $\Delta\Phi_0 \approx 0,82 \text{ В}$. Такая величина характерна для кремниевых p - n переходов. У германиевых переходов потенциальный барьер обычно не превышает $0,4 \text{ В}$ благодаря значительно меньшему собственному сопротивлению.

Из формулы (2-4в) видно, что высота равновесного барьера зависит от температуры, прежде всего, через параметры φ_T и ρ_i . Используя выражения (1-3), (1-36а) и (1-15), можно записать температурный коэффициент в приближенном виде:

$$\frac{d(\Delta\Phi_0)}{dT} = \frac{\Delta\Phi_0 - \varphi_3}{T} < 0, \quad (2-5а)$$

где $\Delta\Phi_0$ соответствует температуре T ; например, для $T = 300 \text{ К}$ нужно подставить $\Delta\Phi_0 (300 \text{ К})$.

Если учесть температурную зависимость подвижностей, входящих в величины ρ_n и ρ_p , получим уточненное значение:

$$\frac{d(\Delta\Phi_0)}{dT} = \frac{\Delta\Phi_0 - (\varphi_3 + c\varphi_T)}{T}, \quad (2-5б)$$

где $c = c_n + c_p$ [см. (1-32)]. Для кремния $c \approx 5$, для германия $c \approx 4$. При $T = 300 \text{ К}$ для кремния $d(\Delta\Phi_0)/dT \approx -1,4 \text{ мВ/}^\circ\text{С}$.

Чтобы вычислить равновесную ширину потенциального барьера l_0 , воспользуемся распределением примесей и зарядов, показанным на рис. 2-7, а и б (оно соответствует структуре перехода на рис. 2-3, в). Примем, что поле в переходе направлено вдоль оси x и отсутствует в поперечных направлениях. Тогда действительно уравнение Пуассона (1-80), где плотности заряда в обеих частях ступенчатого перехода постоянны и определяются концентрациями соответствующих примесей (рис. 2-7, б):

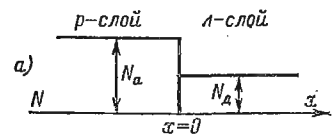
$$\lambda_p = -qN_a; \quad \lambda_n = +qN_d.$$

Соответственно напряженность электрического поля равна нулю вне перехода и изменяется по ломаной линии в пределах перехода (рис. 2-7, в):

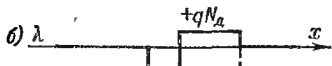
$$E_p = \frac{-qN_a}{\varepsilon_0 \varepsilon} (x + l_p); \quad x \leq 0; \quad (2-6a)$$

$$E_n = \frac{-qN_d}{\varepsilon_0 \varepsilon} (l_n - x); \quad x \geq 0; \quad (2-6б)$$

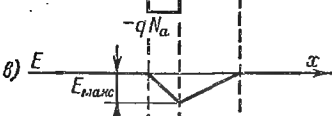
при этом потенциал в пределах перехода будет изменяться по квадратичному закону с точкой перегиба в месте излома кривой $E(x)$ (рис. 2-7, г):



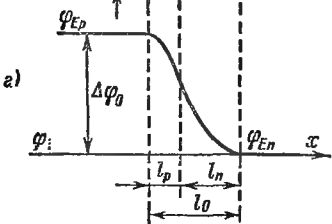
$$\varphi_p - \varphi_{Ep} = -\frac{qN_a}{2\varepsilon_0 \varepsilon} (x + l_p)^2; \quad x \leq 0; \quad (2-7a)$$



$$\varphi_n - \varphi_{En} = \frac{qN_d}{2\varepsilon_0 \varepsilon} (x - l_n)^2; \quad x \geq 0; \quad (2-7б)$$



где φ_{Ep} и φ_{En} — электростатические потенциалы соответствующих слоев вне перехода¹.



Приравнявая $E_p(0)$ и $E_n(0)$, получаем соотношение между шириной перехода в p и n слоях:

$$\frac{l_p}{l_n} = \frac{N_d}{N_a}. \quad (2-8)$$

Рис. 2-7. Распределение концентрации примесей (а), плотности заряда (б), напряженности поля (в) и потенциала (г) в ступенчатом p - n переходе.

Если переход несимметричен и $N_a \gg N_d$, то $l_p \ll l_n$ и, значит, $l_0 \approx l_n$, т. е. переход сосредоточен в высокоомном n -слое. Это обстоятельство отмечалось в предыдущем разделе.

Приравнявая $\varphi_p(0)$ и $\varphi_n(0)$, используя соотношения $l_0 = l_p + l_n$, (2-8) и (2-3), можно получить зависимость между высотой барьера $\Delta\varphi_0$ и шириной перехода l_0 в следующем общем виде:

$$l_0 = \sqrt{\frac{2\varepsilon_0 \varepsilon \Delta\varphi_0}{q} \left(\frac{1}{N_d} + \frac{1}{N_a} \right)}. \quad (2-9a)$$

Для несимметричного перехода при $N_a \gg N_d$ получаем:

$$l_0 = \sqrt{\frac{2\varepsilon_0 \varepsilon \Delta\varphi_0}{qN_d}} \quad (2-9б)$$

[ср. с (1-90б)].

¹ Формулы (2-7) записаны исходя из соотношения $E = d\varphi_E/dx$, т. е. для электронных потенциалов [см. замечание к формулам (1-72)].

Пологая $N_d = n_n$ и выражая концентрацию n_n через удельное сопротивление ρ_n по формуле (1-36б), получаем:

$$l_0 = \sqrt{2\varepsilon_0 \varepsilon_{\mu n} \rho_n \Delta\phi_0}. \quad (2-9в)$$

Например, кремниевый переход с параметрами $\rho_n = 1 \text{ Ом} \cdot \text{см}$ и $\Delta\phi_0 = 0,75 \text{ В}$ имеет ширину $l_0 \approx 5 \cdot 10^{-5} \text{ см}$ (0,5 мкм). С увеличением удельного сопротивления равновесная ширина ступенчатого перехода увеличивается, с уменьшением удельного сопротивления — уменьшается. Однако зависимость $l_0(\rho)$, даваемая формулой (2-9в), не очень точна, если концентрация N_d сравнима с собственной концентрацией n_i , т. е. если $n \neq N_d$, см. (1-24а).

Сравнивая выражения (2-9б) и (1-90а) при $n_0 = N_d$, приходим к выводу, что отношение равновесной ширины несимметричного перехода к дебаевской длине (в высокоомном слое) определяется отношением $2 \Delta\phi_0 / \varphi_T$. Используя (2-5), можно записать:

$$l_0 / l_D = \sqrt{A - 4,6 \lg(\rho_p \rho_n)},$$

где $A_{Si} \approx 52$ и $A_{Ge} \approx 19$. Если $\rho_p = 0,01 \text{ Ом} \cdot \text{см}$ и $\rho_n = 1 \text{ Ом} \cdot \text{см}$, то для кремния $l_0 / l_D \approx 7,8$, а для германия $l_0 / l_D \approx 5,3$. Такие соотношения оправдывают принятую идеализацию структуры перехода (см. с. 91 и рис. 2-3, в). В реальном диапазоне удельных сопротивлений полученные значения l_0 / l_D практически не меняются.

Анализ перехода в неравновесном состоянии. Подключим источник э. д. с. U между p - и n -слоями. Приложенная э. д. с. нарушает равновесие в системе и вызывает протекание тока. При этом высота потенциального барьера должна измениться, так как при равновесном значении $\Delta\phi_0$ потоки носителей через переход уравновешены и тока быть не может.

Выше было показано (см. рис. 2-4), что удельное сопротивление обедненного слоя на несколько порядков выше, чем удельное сопротивление p - и n -слоев диода. Поэтому внешнее напряжение почти полностью падает на переходе, а значит, *изменение высоты потенциального барьера должно быть равно значению приложенной э. д. с.*

Когда э. д. с. U приложена плюсом к p -слою, высота барьера уменьшается¹ (рис. 2-8, а) и становится равной:

$$\Delta\phi = \Delta\phi_0 - U. \quad (2-10)$$

Такое включение перехода называется *прямым*. При отрицательном потенциале на p -слое (рис. 2-8, б) высота барьера увеличивается и в формуле (2-10) нужно изменить знак перед U . Такое включение называется *обратным*.

Изменение высоты барьера с помощью внешнего напряжения приводит к двум главным следствиям: во-первых, изменяется ши-

¹ Положительное приращение потенциала соответствует уменьшению потенциальной энергии электрона, т. е. сдвигу энергетических уровней «вниз».

рина перехода, вторых, изменяются граничные концентрации носителей. Рассмотрим эти зависимости на примере несимметричного перехода с низкоомным p -слоем.

Подставляя значение $\Delta\varphi$ из (2-10) в (2-9б), получаем:

$$l = \sqrt{\frac{2\varepsilon_0\varepsilon(\Delta\varphi_0 - U)}{qN_d}} = l_0 \sqrt{\frac{\Delta\varphi_0 - U}{\Delta\varphi_0}}, \quad (2-11)$$

где l_0 — равновесная ширина потенциального барьера.

Как видим, переход сужается при прямом напряжении ($U > 0$) и расширяется при обратном напряжении ($U < 0$). Однако в первом случае полученное выражение является чисто качественным,

так как погрешность, обусловленная идеализацией перехода (пренебрежением зарядами подвижных носителей), оказывается более существенной, чем в равновесном состоянии. В то же время при обратном напряжении, удовлетворяющем неравенству $|U| \gg \varphi_T$, выражение (2-11) оказывается весьма точным и широко используется на практике. Особенно часто имеет место соотношение $|U| \gg \Delta\varphi_0$, при котором

$$l \approx \sqrt{\frac{2\varepsilon_0\varepsilon}{qN_d}} |U| = l_0 \sqrt{\frac{|U|}{\Delta\varphi_0}}. \quad (2-12)$$

Рис. 2-8. Смещение перехода в прямом (а) и обратном (б) направлениях.

Подставляя в формулы (2-4) высоту барьера (2-10) вместо равновесного значения $\Delta\varphi_0$ и считая концентрации основных носителей n_{n0} и p_{p0} неизменными, получаем для граничных концентраций неосновных носителей выражения:

$$\begin{aligned} p_n &= p_{p0} e^{-\Delta\varphi/\varphi_T} = (p_{p0} e^{-\Delta\varphi_0/\varphi_T}) e^{U/\varphi_T}; \\ n_p &= n_{n0} e^{-\Delta\varphi/\varphi_T} = (n_{n0} e^{-\Delta\varphi_0/\varphi_T}) e^{U/\varphi_T}. \end{aligned}$$

Учитывая, что в скобках стоят равновесные граничные концентрации, определяемые формулами (2-4), запишем полученные выражения в следующей форме:

$$p_n = p_{n0} e^{U/\varphi_T}; \quad (2-13a)$$

$$n_p = n_{p0} e^{U/\varphi_T}. \quad (2-13б)$$

Если напряжение U приложено в прямом направлении, то согласно (2-13) концентрации p_n и n_p на границах перехода возрастают по сравнению с равновесными значениями p_{n0} и n_{p0} . Иначе говоря, в каждом из слоев появляются избыточные неосновные носители, т. е. имеет место инжекция (рис. 2-9, а).

Если напряжение U приложено в обратном направлении, то граничные концентрации p_n и n_p уменьшаются по сравнению с равновесными значениями, т. е. имеет место экстракция (рис. 2-9, б).

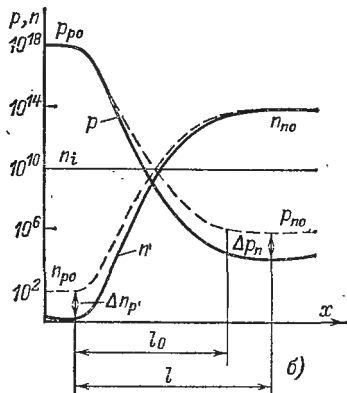
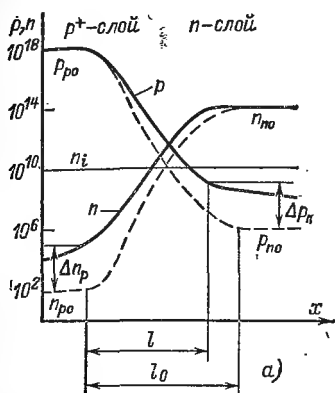


Рис. 2-9. Распределение носителей в кремниевом переходе при прямом (а) и обратном (б) напряжениях. Пунктиром показаны распределения в равновесном состоянии.

Значения избыточных концентраций на границах перехода найдем, вычитая из p_n и n_p соответственно равновесные концентрации p_{n0} и n_{p0} :

$$\Delta p_n = p_{n0}(e^{U/\varphi_T} - 1); \quad (2-14a)$$

$$\Delta n_p = n_{p0}(e^{U/\varphi_T} - 1). \quad (2-14б)$$

Сравним граничные избыточные концентрации в слоях p и n , разделив (2-14а) на (2-14б) и заменив в правой части концентрации p_{n0} и n_{p0} на p_{p0} и n_{n0} по формуле (1-16). Тогда

$$\frac{\Delta p_n}{\Delta n_p} = \frac{p_{p0}}{n_{n0}}. \quad (2-15)$$

В несимметричных переходах концентрации p_{p0} и n_{n0} сильно различаются, поэтому концентрация инжектированных неосновных носителей будет гораздо больше в высокоомном слое, чем в низкоомном. Таким образом, в несимметричных переходах инжекция имеет односторонний характер: неосновные носители инжектируются в основном из низкоомного слоя в высокоомный.

Инжектирующий слой с относительно малым удельным сопротивлением называют *эмиттером*, а слой с относительно большим удельным сопротивлением, в который инжектируются неосновные для него носители, — *базой*.

Формулы (2-13) и (2-14) подтверждают, что в режиме экстракции граничные концентрации неосновных носителей могут быть сколь

удовно малы, но всегда положительно значительны, а избыточные концентрации отрицательны, но по модулю всегда меньше равновесных значений (см. с. 85, п. 2). Обратные напряжения, при которых $p_n \ll p_{n0}$, $n_p \ll n_{p0}$ и соответственно $\Delta p_n \approx -p_{n0}$, $\Delta n_p \approx -n_{p0}$, определяются условием

$$|U| > (3-4) \varphi_T. \quad (2-16)$$

В дальнейшем запись $|U| > (3-4) \varphi_T$ будет пониматься именно в таком смысле. Заметим, что условие (2-16) применительно к прямым напряжениям означает $p_n \gg p_{n0}$ и $n_p \gg n_{p0}$, т. е. позволяет пренебречь единицей в выражениях (2-14).

В теории полупроводниковых диодов и транзисторов зависимости (2-13), (2-14) играют весьма важную роль, поэтому полезно исследовать границы их применимости. Для этого следует напомнить, что указанные зависимости получены из (2-4) путем простой замены равновесной величины $\Delta \varphi_0$ на неравновесную величину $\Delta \varphi = \Delta \varphi_0 - U$, а также в предположении неизменных концентраций основных носителей n_{n0} и p_{p0} . Последнее предположение соответствует условию низкого уровня инжекции в базе [см. (1-109)]¹. Что касается замены $\Delta \varphi_0$ на $\Delta \varphi$, то в ее основе лежит понятие квазиравновесного (почти равновесного) состояния перехода при наличии внешнего напряжения. Действительно, поскольку выражения (2-4) получены из условия бальцмановского равновесия (т. е. равенства диффузионных и дрейфовых составляющих токов в переходе), то использование этих же выражений для неравновесного состояния означает, что равновесие нарушено несущественно, т. е. результирующий (разностный) ток значительно меньше каждой из составляющих — диффузионной и дрейфовой: $j \ll j_{\text{диф}} \approx j_{\text{др}}$.

Для того чтобы конкретизировать это условие квазиравновесия, нужно оценить одну из равновесных составляющих, например диффузионную. Строгая оценка затрудняется тем, что внутри перехода градиенты концентрации существенно меняются (рис. 2-3, б и 2-4). Поэтому ограничимся грубой оценкой применительно к дырочным токам, которые, как уже отмечалось, играют главную роль в несимметричном переходе с p^+ -эмиттером.

Средний градиент концентрации дырок в переходе примем равным $[p_{p0} - p_n(0)]/l_0$, где $p_n(0) = p_{n0} + \Delta p_n(0) \approx \Delta p_n(0)$ (условие $\Delta p_n(0) \gg p_{n0}$ практически всегда выполняется). Этим градиентом согласно (1-73а) определяется составляющая $j_{\text{диф}}$ в переходе. Что касается результирующего тока j , то он согласно (1-116) определяется градиентом $\Delta p_n(0)/L_p$, где L_p — диффузионная длина дырок в базе ($L_p \gg l_0$). Тогда из неравенства $j \ll j_{\text{диф}}$ легко получить условие квазиравновесия p^+ - n перехода:

$$\Delta p_n(0) \ll p_{p0} \frac{L_p}{L_p + l_0} \approx p_{p0}. \quad (2-17)$$

На зонных диаграммах состояние квазиравновесия в переходе отражается почти горизонтальным расположением квазиуровней Ферми для дырок и

¹ Уровень инжекции в эмиттере, как следует из (2-15), всегда значительно ниже, чем в базе. Поэтому в большинстве реальных случаев можно считать $p_p = p_{p0} = \text{const}$ для любых токов, тогда как концентрация n_n с ростом тока может немного превысить равновесное значение n_{n0} , если уровень инжекции в базе много больше единицы.

электронов (рис. 2-10). В самом деле, из определений (1-11) следует:

$$\frac{d\varphi_{F_{n,p}}}{dx} = \frac{d\varphi_{c,v}}{dx} \pm \frac{d\chi_{n,p}}{dx}.$$

Поскольку градиенты в этом соотношении пропорциональны соответствующим потокам и поскольку результирующие потоки, отражаемые левой частью, по условию много меньше каждой из составляющих, отражаемых правой частью, можно положить в н у т р и перехода $d\varphi_{F_{n,p}}/dx = 0$. Вне перехода квазиуровни Ферми спадают до равновесных значений в соответствующем слое. Градиенты этих спадов характеризуют некомпенсированные диффузионные потоки н е о с н о в н ы х носителей на протяжении нескольких диффузионных длин. Далее,

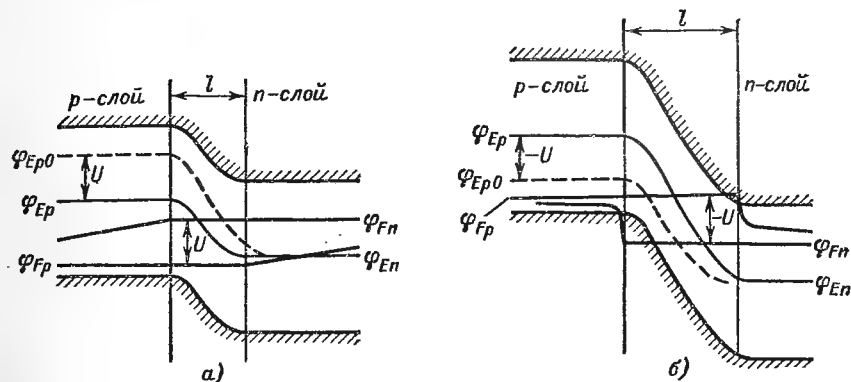


Рис. 2-10. Квазиуровни Ферми при прямом (а) и обратном (б) напряжениях на переходе (считается $\varphi_{E_n} = \text{const}$). Пунктиром показан электростатический потенциал в равновесном состоянии.

после слияния, квазиуровни Ферми имеют очень небольшой постоянный наклон (градиент потенциала), характеризующий наличие дрейфового тока и омического поля в однородных слоях.

Из рис. 2-9 видно, что при инжекции и экстракции концентрации электронов и дырок в переходе изменяются в одном и том же направлении, т. е. произведение концентраций не может оставаться постоянным, как в случае равновесия [см. (1-16)]. Чтобы оценить произведение np в условиях квазиравновесия, заменим в выражениях (1-17) единый уровень φ_F соответственно на φ_{Fn} и φ_{Fp} и учтем, что $\varphi_{Fn} - \varphi_{Fp} = U$ (см. рис. 2-10). Тогда получаем:

$$np = n_i^2 e^{U/\varphi_T}. \quad (2-18)$$

Из этого соотношения следует, что *прямое смещение перехода эквивалентно увеличению собственной концентрации в нем и, следовательно, сопровождается уменьшением удельного сопротивления обедненного слоя. Обратное смещение, напротив, эквивалентно уменьшению собственной концентрации в переходе*, причем, если условие (2-16) выполняется достаточно сильно (например, для кремния, если $|U| > 50\varphi_T$), то концентрации носителей в обедненном слое падают буквально до нескольких единиц в кубическом санти-

метре, т. е. переход точно соответствует идеализации, принятой на рис. 2-3, в. Именно этим объясняется высокая точность формулы (2-11) при обратных напряжениях.

Плавные p - n переходы. Плавный переход образуется «контактом» двух слоев p и n , из которых хотя бы один неоднородный. На практике в неоднородных слоях примесь распределена либо по закону функции ошибок (см. сноску на с. 84), либо по экспоненциальному закону (1-93). Однако для анализа обычно принимают линейное распределение примеси в пределах перехода (рис. 2-11).

Анализ плавного перехода осложняется наличием внутренних электрических полей в неоднородных полупроводниках (см. рис. 1-17, а и § 1-12). Однако, учитывая, что эти поля, как правило, в десятки раз слабее, чем поле в переходе, упростим задачу и будем считать слои, примыкающие к переходу, квазиоднородными. Тогда высоту равновесного потенциального барьера можно определить по формулам (2-4), подставляя в них концентрации носителей на тех участках, которые непосредственно прилегают к переходу.

Для определения ширины перехода снова примем, что в области перехода нет свободных носителей и что, следовательно, эта область резко ограничена, а пространственный заряд в ней создается только ионами примесей. Тогда плотность заряда в переходе λ будет изменяться по ломаной линии (рис. 2-11, б) с разными градиентами на разных участках: qN'_a , $q(N'_a + N'_d)$ и qN'_d , где N'_a и N'_d — градиенты концентрации акцепторов и доноров.

Если переход значительно шире области перекрытия ($l \gg d$), то средним участком d можно пренебречь и считать, что диаграмма λ состоит только из двух участков с градиентами плотности заряда qN'_a и qN'_d . В этом случае, считая, что плотность заряда линейно меняется от $\lambda = -qN'_a l_p$ (на левой границе перехода) до нуля и затем от нуля до $\lambda = +qN'_d l_n$ (на правой границе перехода), напряженность поля нетрудно получить из уравнения (1-80):

$$E_p = -\frac{qN'_a}{2\epsilon_0 e} (l_p^2 - x^2), \quad x \leq 0;$$

$$E_n = -\frac{qN'_d}{2\epsilon_0 e} (l_n^2 - x^2), \quad x \geq 0.$$

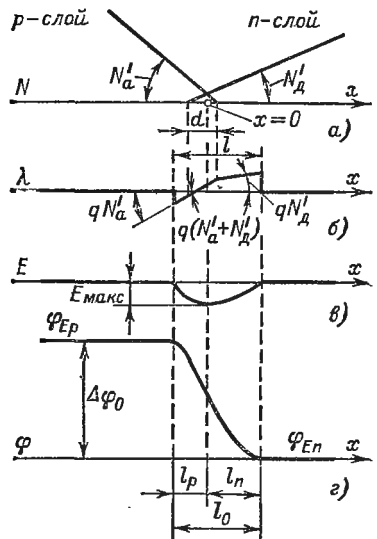


Рис. 2-11. Распределение концентрации примесей (а), плотности заряда (б), напряженности поля (в) и потенциала (г) в плавном p - n переходе.

Кривая $E(x)$ показана на рис. 2-11, в в сочлененных квадратных параболах. Соответственно для электростатического потенциала получится кубическая зависимость от координаты (рис. 2-11, з):

$$\varphi_p - \varphi_{E_p} = \frac{-qN'_a}{6\epsilon_0 e} (2l_p^3 - x^3 + 3l_p^2 x), \quad x \leq 0; \quad (2-19a)$$

$$\varphi_n - \varphi_{E_n} = \frac{qN'_d}{6\epsilon_0 e} (2l_n^3 + x^3 - 3l_n^2 x), \quad x \geq 0, \quad (2-19b)$$

где φ_{E_p} и φ_{E_n} — электростатические потенциалы в глубине слоев.

Приравнявая $E_p(0)$ и $E_n(0)$, находим соотношение между шириной перехода в p - и n -слоях:

$$\frac{l_p}{l_n} = \sqrt{\frac{N'_d}{N'_a}}. \quad (2-20)$$

В случае равных градиентов концентрации получаем $l_p = l_n$, т. е. переход симметричен. В случае резко различных градиентов переход сосредоточен в слое с малым градиентом.

Приравнявая $\varphi_p(0)$ и $\varphi_n(0)$ и используя соотношения $l_0 = l_n + l_p$, (2-20) и (2-3), находим ширину равновесного перехода в общем виде:

$$l_0 = \sqrt[3]{\frac{3\epsilon_0 e \Delta\varphi_0}{q} \left(\frac{1}{N'_a} + \frac{1}{\sqrt{N'_a N'_d}} + \frac{1}{N'_d} \right)}. \quad (2-21a)$$

В случае несимметричного перехода, например, если $N'_a \gg N'_d$, выражение (2-21a) упрощается:

$$l_0 = \sqrt[3]{\frac{3\epsilon_0 e \Delta\varphi_0}{qN'_d}}. \quad (2-21b)$$

Эта формула, как и (2-9б), приближенная и для прямого смещения (когда $\Delta\varphi_0$ заменяется на $\Delta\varphi = \Delta\varphi_0 - U$) дает большую погрешность. Однако при обратном напряжении, удовлетворяющем условию $|U| \gg \Delta\varphi_0$, получается формула, аналогичная (2-12):

$$l \approx \sqrt[3]{\frac{3\epsilon_0 e}{qN'_d} |U|} = l_0 \sqrt[3]{\frac{|U|}{\Delta\varphi_0}}, \quad (2-22)$$

точность которой вполне достаточна для практических целей.

Обычно плавный переход является «узким», т. е. можно считать $l < d$ (см. рис. 2-11). Этот случай охватывается выведенными формулами: достаточно заменить в них градиенты N'_a и N'_d суммарным градиентом $N' = N'_a + N'_d$. Такие переходы согласно (2-20) расположены симметрично в обоих слоях: $l_p = l_n$. Их равновесная ширина в общем случае согласно (2-21a) имеет вид:

$$l_0 = \sqrt[3]{\frac{9\epsilon_0 e \Delta\varphi_0}{qN'}}. \quad (2-23a)$$

Для обратных напряжений, удовлетворяющих условию $|U| \gg \Delta\varphi_0$,

$$l \approx \sqrt[3]{\frac{9\epsilon_0 e}{qN'} |U|} = l_0 \sqrt[3]{\frac{|U|}{\Delta\varphi_0}}. \quad (2-23b)$$

В условиях квазиравновесия плавного перехода остаются в силе формулы инжекции и экстракции (2-13) и (2-14).

Односторонний p - n переходы. В случае резкой несимметричности p - n переходов, когда концентрации основных носителей в обоих слоях различаются на 1—2 порядка и более, структура перехода качественно меняется: обедненные слои, расположенные по обе

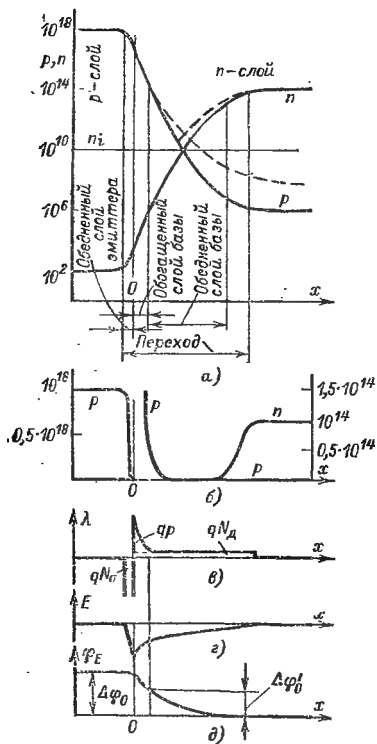


Рис. 2-12. Односторонний кремниевый переход. Распределение носителей в полулогарифмическом (а) и линейном (б) масштабе. Распределение объемных зарядов (в), поля (г) и потенциала (д). Пунктирные линии соответствуют прямому смещению.

стороны от металлургической границы, оказываются разделенными обогащенным слоем, расположенным в высокоомной части перехода (рис. 2-12)¹. Наличие обогащенного слоя не может не сказаться на распределении зарядов, поля и потенциала.

Будем, как и раньше, рассматривать p^+ - n переход, у которого потенциальный барьер сосредоточен в базовой области. Поскольку часть напряжения $\Delta\varphi_0$ падает в обогащенном слое, на долю обедненного слоя приходится напряжение $\Delta\varphi'_0 < \Delta\varphi_0$, которое следует подставлять в формулу (2-96), чтобы найти ширину l_0 (эмиттерный и обогащенный слои значительно тоньше и не влияют на общую ширину перехода). Напряжение $\Delta\varphi'_0$ можно найти из выражения (2-46), заменяя p_{E0} на n_{n0} (так как на условной границе между обогащенным и обедненным слоями в базе $p = n_{n0}$). Далее, выражая p_{n0} через n_{n0} с помощью соотношения (1-16), получаем [27]:

$$\Delta\varphi'_0 = 2\varphi_T \ln \frac{n_{n0}}{n_i} = 2(\varphi_F - \varphi_{En}), \quad (2-24)$$

где вторая форма записи основана на соотношении (1-18а).

Полная высота равновесного потенциального барьера $\Delta\varphi_0$ не зависит от структуры перехода и для односторонних переходов определяется общими формулами (2-4). Остаются в силе и общие формулы инжекции и экстракции (2-13), (2-14)².

¹ На рис. 2-12 показан ступенчатый односторонний переход. Плавные односторонние переходы сложнее для анализа, но основные выводы действительны и для них.

² Наличие обогащенного слоя практически не влияет на распределение внешнего приложенного напряжения; оно по-прежнему почти полностью падает на обедненном слое базы.

Критерием, характеризующим односторонность $p^{++}n$ перехода, т. е. наличие обогащенного слоя, может служить неравенство $p(0) \geq N_d$, где координата $x = 0$ соответствует металлургической границе. Анализ показывает, что критерий односторонности перехода можно записать в виде

$$(z-1) \ln z \geq 2 \ln y, \quad (2-25a)$$

где $z = N_a/N_d$ и $y = N_d/n_i$. Зависимость $z_{\min}(\lg y)$ можно получить путем графического решения (2-25a); тогда (2-25a) принимает вид:

$$z \geq z_{\min} = 3 + 1,5 \lg y. \quad (2-25b)$$

Если $y = 10-10^7$, то $z_{\min}$ лежит в пределах 5—14. Следовательно, односторонний переход получается при соотношении $N_a > (5-14) N_d$, которое всегда выполняется на практике, за исключением некоторых специальных случаев. Таким образом, все реальные несимметричные переходы являются односторонними.

Эмиттерная часть объемного заряда оказывается у односторонних переходов очень узкой. Если оценить ее ширину с помощью соотношения (2-8), полагая $l_n = l_0$, то она, как правило, получается меньше дебаевской длины в эмиттере, что противоречит законам распределения поля в полупроводниках (см. раздел «Эффект поля» в § 1-12). Это значит, что принятая ранее идеализация эмиттерной части перехода (постоянство плотности заряда на участке l_p , рис. 2-3, в) для односторонних переходов неприемлема. На самом деле объемный заряд и поле в эмиттере имеют протяженность порядка дебаевской длины (0,005—0,01 мкм при $N_a = 10^{18} \text{ см}^{-3}$), причем концентрация дырок на этом участке спадает плавно, так что собственно обедненный слой в эмиттере отсутствует. Последнее обстоятельство объясняется тем, что эмиттеры в односторонних переходах представляют собой полуметаллы, которым, как и металлам, обедненные слои несвойственны.

2-3. СПЕЦИАЛЬНЫЕ ТИПЫ ПЕРЕХОДОВ

Помимо электронно-дырочных переходов, рассмотренных выше, в полупроводниковых приборах встречаются и другие типы существенных неоднородностей, которые тоже можно отнести к переходам.

Переходы между примесными и собственными полупроводниками. Предположим, что в одностороннем $p^{++}n$ переходе (рис. 2-12) концентрация доноров в базе непрерывно уменьшается. Тогда обогащенный дырочный слой, прилегающий к металлургической границе, будет расширяться и обогащаться дырками, а роль донорных ионов будет становиться все меньше. В пределе, при $N_d = 0$, образуется $p-i$ переход, у которого обедненный слой в базе отсутствует, а положительный объемный заряд обусловлен только дырками (рис. 2-13).

Распределение поля и потенциала в $p-i$ переходе сложнее, чем в $p-n$ переходах, так как плотность объемного заряда дырок в базе даже приближенно нельзя считать постоянной. Тем не менее

высота равновесного потенциального барьера по-прежнему определяется формулой (2-4б), если вместо p_{n0} подставить n_i . При этом значение $\Delta\phi_0$ оказывается, конечно, меньше, чем в p^+-n переходе.

Отсутствие обедненного слоя в $p-i$ переходе приводит к тому, что внешнее напряжение в значительной мере падает на высоко-

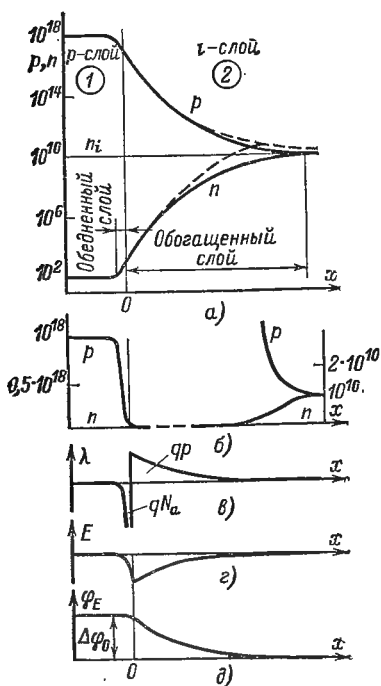


Рис. 2-13. Переход $p-i$. Распределение носителей в полулогарифмическом (а) и линейном (б) масштабе. Распределение объемных зарядов (в), поля (г) и потенциала (д). Пунктирные линии соответствуют прямому смещению.

мере абстракциями, поскольку идеальных собственных полупроводников не существует. Однако эти переходы как бы прокладывают мост между хорошо изученными выше $p-n$ переходами и переходами p^+-p и n^+-n типа, которые часто встречаются на практике.

Рассмотрим переход p^+-p типа, образованный контактом двух дырочных полупроводников с разными концентрациями примесей. Пусть сначала имелся $p-i$ переход (рис. 2-13), в котором слои p и i обозначим 1 и 2. Пусть теперь в собственный слой 2 вводится

нейтральном слое базы, обладающем собственной проводимостью, и лишь часть внешнего напряжения приходится на область перехода и модулирует высоту потенциального барьера¹. Следовательно, нелинейность в $p-i$ структуре выражена слабее, чем в классическом $p-n$ переходе.

Ширина обогащенного слоя в базовой части перехода составляет $(3 \div 4) l_{Di}$. Ширина эмиттерной части, как и в одностороннем переходе, определяется дебаевской длиной l_{Dp} в высоколегированном слое эмиттера (см. § 2-2).

Из рис. 2-13 видно, что прямое смещение (пунктирные линии) приводит к некоторому уменьшению ширины перехода в его высокоомной части и что концентрация электронов меняется относительно сильнее, чем концентрация дырок. Однако поскольку роль электронов в $p-i$ переходе невелика, это не приводит к существенному увеличению удельной проводимости обогащенного слоя.

Переходы между однотипными полупроводниками. Переходы $p-i$ и $n-i$ типа являются в известной

¹ Этот вывод подразумевает достаточную длину нейтральной базы, в противном случае ее сопротивление не будет определяющим.

все большее число акцепторных атомов. Тогда общая структура p - i перехода сохраняется, но по мере увеличения концентрации N_{a2} потенциальный барьер становится все ниже. Это следует из формулы (2-46), где концентрацию p_{n0} нужно заменить концентрацией p_{p2} , значение которой приближается к значению p_{p1} . Механизм понижения барьера заключается в том, что положительный заряд избыточных дырок в слое 2 постепенно уменьшается, так как убывает разность концентраций $p_{p1} - p_{p2}$, вызывающая диффузию дырок из слоя 1 в слой 2. Вместе с уменьшением заряда избыточных дырок, конечно, уменьшается и равный ему отрицательный заряд «обнаженных» акцепторов в слое 1.

По мере увеличения концентрации N_{a2} уменьшается также ширина перехода, поскольку протяженность обогащенного слоя справа от металлургической границы (рис. 2-14) определяется дебаевской длиной l_{Dp2} , а она обратно пропорциональна величине $\sqrt{N_{a2}}$. Таким образом, для p^+-p перехода характерны малая высота потенциального барьера (несколько ϕ_T) и малая ширина его. При равенстве $N_{a2} = N_{a1}$ переход исчезает и получается однородный полупроводник.

Спецификой p^+-p и n^+-n переходов (как и p - i перехода) является отсутствие инъекции и экстракции основных носителей при наличии смещения, поскольку основные носители в обоих слоях одинаковы.

Основная часть приложенного напряжения падает на нейтральной части высокоомного p -слоя, а модуляция проводимости перехода малосущественна. Соответственно вентильные свойства практически отсутствуют, и система является просто неоднородным полупроводником. Такой вариант используется в омических контактах (см. ниже). Подробный анализ переходов между однотипными полупроводниками можно найти в [28].

2-4. КОНТАКТЫ МЕТАЛЛ-ПОЛУПРОВОДНИК

Контакты полупроводника с металлом [5, 27] играют важную роль в полупроводниковых приборах. Структура и свойства этих контактов зависят в первую очередь от взаимного расположения

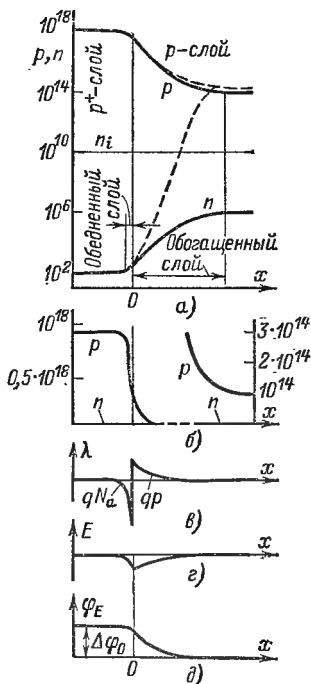


Рис. 2-14. Переход p^+-p . Распределение носителей в полупроводнике (а) и линейном (б) масштабе. Распределение объемных зарядов (в), поля (г) и потенциала (д). Пунктирные линии соответствуют прямому смещению.

уровней Ферми в исходных слоях. На рис. 2-15—2-17 сверху показаны зонные диаграммы разделенных слоев, а внизу — зонные диаграммы соответствующих контактов. Поскольку уровень Ферми в металле всегда расположен в зоне проводимости и для равновесной системы должен быть единым, нетрудно в каждом случае оценить характер результирующей диаграммы.

Выпрямляющие контакты. На рис. 2-15, а имеем $\varphi_{Fm} > \varphi_{Fr}$. Это значит, что энергетические уровни, соответствующие зоне

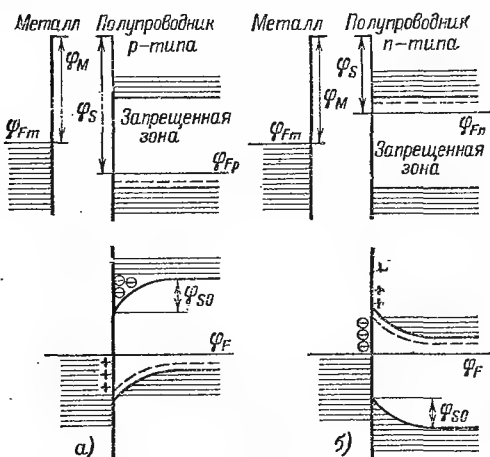


Рис. 2-15. Зонные диаграммы выпрямляющих контактов металла с полупроводником.

а — контакт с полупроводником *p*-типа; б — контакт с полупроводником *n*-типа.

электроны переходят из полупроводника в металл и соответственно уровни искривляются вверх. Область искривления зон (т. е. область пространственных зарядов) имеет протяженность, примерно равную дебаевской длине (1-90а), которая в зависимости от удельного сопротивления полупроводника может составлять десятки, сотые доли микрона и меньше (вплоть до единиц нанометров).

Описанный выше обмен электронами между металлом и полупроводником обычно характеризуют не разностью «исходных» уровней Ферми, а разностью *работ выхода*. Работой выхода электрона из твердого тела называют энергетическое «расстояние» между уровнем свободного электрона вне твердого тела и уровнем Ферми (см. рис. 2-15, где работы выхода из металла и полупроводника обозначены соответственно через φ_m и φ_s). Разность работ выхода

проводимости полупроводника, заполнены в металле больше, чем в полупроводнике. Следовательно, после соприкосновения слоев часть электронов перейдет из металла в полупроводник и создаст отрицательный заряд на границе с металлом¹. Наличие дополнительных электронов согласно (1-10а) приводит к уменьшению расстояния между уровнем Ферми и дном зоны проводимости в этой области, поэтому энергетические уровни полупроводника искривляются вниз. На рис. 2-15, б показан обратный случай, когда после соприкосновения слоев

¹ Практически создать качественный контакт между металлом и полупроводником путем их простого соприкосновения невозможно (см. аналогичное замечание по поводу *p-n* перехода на с. 88). Реальные контакты такого рода в настоящее время создаются напылением металла на полупроводник в вакууме.

Контактная разность потенциалов φ_{MS} между металлами и кремнием

$N, \text{см}^{-3}$	Al		Au	
	<i>n</i> -Si	<i>p</i> -Si	<i>n</i> -Si	<i>p</i> -Si
10^{14}	-0,36	-0,82	+0,54	+0,08
10^{15}	-0,30	-0,88	+0,60	+0,02
10^{16}	-0,24	-0,94	+0,66	-0,04
10^{17}	-0,18	-1,00	+0,72	-0,10

$\varphi_{MS} = \varphi_M - \varphi_S$, выраженную в вольтах, называют *контактной разностью потенциалов* (см. табл. 2-1).

В зависимости от соотношения значений φ_M и φ_S электроны при сближении слоев переходят либо из металла в полупроводник (случай $\varphi_{MS} < 0$, рис. 2-15, а), либо из полупроводника в металл (случай $\varphi_{MS} > 0$, рис. 2-15, б). В обоих случаях происходит искривление энергетических зон и на границе раздела появляется тот равновесный поверхностный потенциал φ_{s0} , о котором упоминалось на с. 70. При непосредственном контакте металла с полупроводником можно считать $\varphi_{s0} = \varphi_{MS}$. Если же слои разделены диэлектриком (в частности, вакуумом или воздухом), то часть напряжения φ_{MS} падает на диэлектрике и тогда $\varphi_{s0} < \varphi_{MS}$. Что касается напряжения спрямления зон U_F (см. с. 70), то оно всегда равно напряжению φ_{MS} независимо от наличия или отсутствия диэлектрика¹.

Оба контакта, показанных на рис. 2-15, а и б, характеризуются тем, что концентрация основных носителей в приконтактном граничном слое полупроводника понижена по сравнению с концентрацией их вдали от контакта. Следовательно, *граничный слой обладает повышенным удельным сопротивлением и поэтому определяет сопротивление всей системы*. В зависимости от полярности приложенного напряжения меняются высота приповерхностного потенциального барьера и соответственно сопротивление граничного слоя. Так, если внешнее напряжение приложено плюсом к металлу и минусом к полупроводнику, то потенциальный барьер на рис. 2-15, а повышается, а на рис. 2-15, б понижается. При этом граничный слой на рис. 2-15, а еще больше обедняется дырками и будет иметь повышенное сопротивление, а граничный слой на рис. 2-15, б обогащается электронами и будет иметь пониженное сопротивление по сравнению с равновесным состоянием. Значит, такая полярность будет обратной для перехода на рис. 2-15, а и прямой для перехода на рис. 2-15, б. Если изменить полярность приложенного напряжения, то по аналогичным причинам сопротивление граничного слоя на рис. 2-15, а пони-

¹ Разумеется, этот вывод предполагает, что разность работ выхода является единственной причиной искривления зон.

зится и переход будет работать в прямом направлении, а сопротивление граничного слоя на рис. 2-15, б повысится и этот переход будет работать в обратном направлении. Таким образом, контакты на рис. 2-15 хотя и не обеспечивают инъекции неосновных носителей, но обладают вентильными свойствами. Такие контакты лежат в основе диодов Шоттки (см. § 3-4).

Особый интерес представляет контакт, показанный на рис. 2-16. Его особенность состоит в том, что уровень Ферми металла в исходном состоянии лежит ниже середины запрещенной зоны полупроводника n -типа.

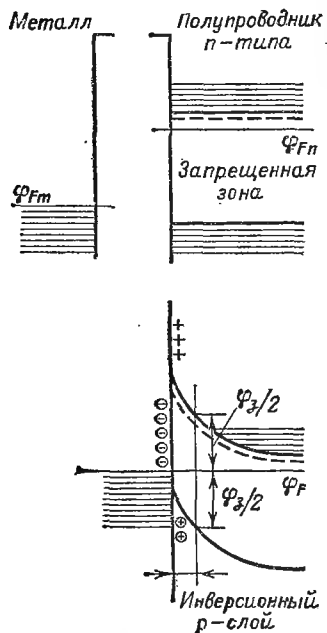


Рис. 2-16. Зонная диаграмма контакта, при котором образуется инверсионный слой.

Такое расположение, как известно, характерно для дырочных полупроводников (см. § 1-8). Следовательно, в данном случае вблизи поверхности полупроводника n -типа образовался тонкий слой полупроводника с обратным типом проводимости (p -типа). Этот слой называют *инверсионным* (см. раздел «Эффект поля», § 1-12). В целом, как видим, получился p - n переход, полностью расположенный внутри исходной пластинки полупроводника. С физической точки зрения образование инверсионного слоя объясняется тем, что электронов в зоне проводимости полупроводника (в его граничном слое) оказывается недостаточно для равновесия системы и в металл должно перейти некоторое количество электронов из валентной зоны; в результате, естественно, образуются дырки. Толщина инверсионного слоя («канала»), как уже отмечалось (с. 71), обычно составляет 0,001—0,002 мкм.

Невыпрямляющие (омические) контакты. Пусть для контакта металла с полупроводником p -типа имеет место соотношение $\varphi_{Fm} < \varphi_{Fp}$, а для контакта металла с полупроводником n -типа — соотношение $\varphi_{Fm} > \varphi_{Fn}$. В этих случаях искривление зон в полупроводниках получается обратным по сравнению с тем, какое показано на рис. 2-15, а и б, т. е. граничные слои оказываются не обедненными, а обогащенными основными носителями (рис. 2-17). Соответственно удельные сопротивления граничных слоев оказываются значительно меньше, чем основных, нейтральных слоев

Невыпрямляющие (омические) контакты. Пусть для контакта металла с полупроводником p -типа имеет место соотношение $\varphi_{Fm} < \varphi_{Fp}$, а для контакта металла с полупроводником n -типа — соотношение $\varphi_{Fm} > \varphi_{Fn}$. В этих случаях искривление зон в полупроводниках получается обратным по сравнению с тем, какое показано на рис. 2-15, а и б, т. е. граничные слои оказываются не обедненными, а обогащенными основными носителями (рис. 2-17). Соответственно удельные сопротивления граничных слоев оказываются значительно меньше, чем основных, нейтральных слоев

полупроводника (вдали от границы), так что наличие граничного слоя в системе оказывается малосущественным с точки зрения ее суммарного сопротивления. Последнее будет близко к сопротивлению нейтрального слоя полупроводника и, следовательно, почти не будет зависеть от полярности и величины внешнего напряжения. Такие невыпрямляющие переходы являются основой омических контактов.

В терминах контактной разности потенциалов случай на рис. 2-17, а соответствует полярности $\varphi_{MS} > 0$ (т. е. переходу электронов из полупроводника), а случай на рис. 2-17, б — полярности $\varphi_{MS} < 0$ (т. е. переходу электронов из металла). В обоих случаях контактные разности потенциалов приняты небольшими, поэтому искривления зон на рис. 2-17 сравнительно малы и уровень Ферми не пересекает разрешенные зоны в полупроводнике. При больших значениях $|\varphi_{MS}|$ искривления зон будут больше и уровень Ферми будет частично проходить либо через валентную зону дырочного полупроводника (рис. 2-17, а), либо через зону проводимости электронного полупроводника (рис. 2-17, б). Это значит, что соответствующие граничные участки полупроводника превращаются в полуметаллы (см. § 1-4, 1-6).

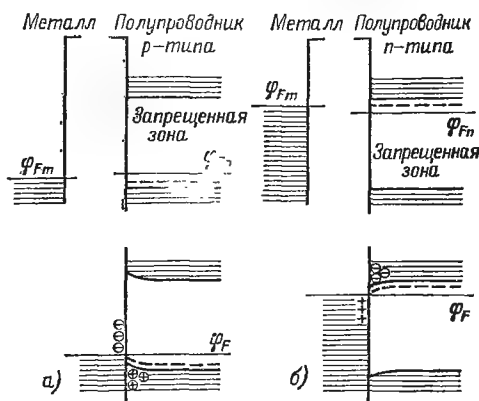


Рис. 2-17. Зонные диаграммы невыпрямляющих контактов металла с полупроводником.

а — контакт с полупроводником *p*-типа; б — контакт с полупроводником *n*-типа.

Омические контакты осуществляют в местах присоединения внешних выводов к полупроводниковому слою. Такие контакты не образуют дополнительного («паразитного») перехода. Получение омических контактов является задачей не менее важной, чем получение рабочих *p-n* переходов. Типичная структура омического контакта показана на рис. 2-18. Как видим, эта структура состоит из двух переходов: n^+-n и $m-n^+$, где через *m* обозначен слой металла. Оба перехода не являются инжектирующими, как было показано в предыдущих разделах. Кроме того, они не обладают и вентильными свойствами. Поэтому в целом структура $n-n^+-m$ ведет себя почти как омическое сопротивление слоя *n* при любой полярности напряжения. Рассмотрим механизм прохождения токов. Пусть напряжение приложено минусом к слою *n* и плюсом к металлу. Тогда потенциалы слоев *n* и n^+ повысятся, высота барьера $n-n^+$ увеличится, а высота барьера n^+-m уменьшится (рис. 2-18, б). Электроны из *n*-слоя будут свободно переходить в n^+ -слой независимо от высоты барьера $n-n^+$, а понижение барьера n^+-m обеспечит переход электронов из n^+ -слоя в *m*-слой. Пусть теперь напряжение приложено плюсом к *n*-слою. При этом потенциалы *n*- и n^+ -слоев пони-

зятся, и высота барьера n - n^+ делается меньше; соответственно электроны n^+ -слоя смогут переходить в n -слой. Барьер n^+ - m повысится, но так как он очень тонкий¹, то электроны слоя m будут проходить его за счет туннельного эффекта [29, § 24], как показано на рис. 2-18, в.

Таким образом, важнейшим свойством омического контакта является его двусторонняя проводимость. Другое важное свойство

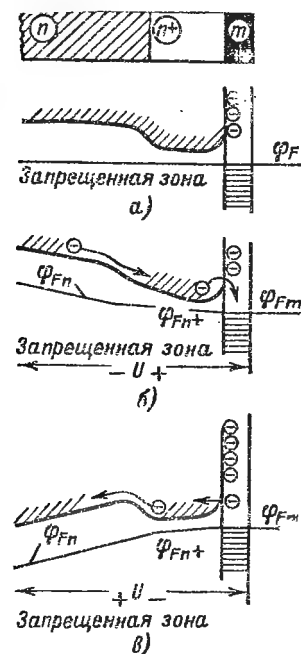


Рис. 2-18. Зонные диаграммы омического контакта.

а — равновесное состояние; б — обратное напряжение; в — прямое напряжение.

контактов находит алюминий. Он легко напыляется на поверхность кремния и далее при достаточно высокой температуре сплавляется с ним. Если кремний имеет проводимость p -типа, то алюминий, будучи акцептором, способствует повышению проводимости приконтактного слоя. Если же кремний имеет проводимость n -типа, то при сравнительно небольшой концентрации доноров

¹ Например, при $n^+ = N_A^+ = 10^{18} \text{ см}^{-3}$ получается дебаевская длина $l_D \approx 50 \cdot 10^{-8} \text{ см}$ (0,005 мкм).

² Более общим граничным условием для омического контакта является равенство концентрации носителей сумме равновесной концентрации и члена, обратно пропорционального скорости рекомбинации.

акцепторные атомы алюминия могут создать паразитный приконтактный p -слой; в результате вместо омического контакта получится выпрямляющий p - n переход. Чтобы избежать такого осложнения, поверхность n -кремния дополнительно легируют донорами, превращая ее в n^+ -слой, после чего контакт приобретает структуру, показанную на рис. 2-18.

В заключение заметим, что приповерхностные потенциальные барьеры могут быть обусловлены не только контактом с металлом, как показано выше, или внешним электрическим полем (см. раздел «Эффект поля» в гл. 1), но и чисто «внутренними» причинами: наличием *поверхностных энергетических уровней* или, как говорят, *поверхностных состояний* (с. 46). Эти уровни в принципе могут иметь как донорный, так и акцепторный характер, и тогда приповерхностная область полупроводника обогащена соответственно либо электронами, либо дырками. В кремнии — основном материале современной транзисторной техники — поверхностные уровни являются донорными. Значит, энергетические зоны в кремнии «с самого начала» (т. е. до контакта с металлом и до подачи внешнего напряжения) искривлены «вниз» и на поверхности имеется начальный поверхностный потенциал ϕ_{s0} отрицательной полярности. Этот начальный потенциал меняется в ту или иную сторону в зависимости от знаков контактной разности потенциалов «металл—полупроводник» и приложенного напряжения.

2-5. АНАЛИЗ ИДЕАЛИЗИРОВАННОГО ДИОДА

Несмотря на то что диод представляет собой один из простейших полупроводниковых приборов, процессы, происходящие в нем, достаточно сложны. Для того чтобы выяснить главные особенности диода, проведем сначала упрощенный анализ, а затем (в последующих параграфах) уточним полученные результаты.

Исходные предпосылки. Будем, как и раньше, считать p - n переход несимметричным и p -слой значительно больше легированным, чем n -слой. При этом, как известно, инжекция и экстракция носят односторонний характер [см. (2-15)]; следовательно, можно сосредоточить внимание на анализе процессов в базе, а результаты анализа распространить затем на аналогичные, но менее существенные процессы в эмиттере.

Анализ существенно упрощается, если принять следующие допущения.

1. Слой базы является ярко выраженным электронным полупроводником. Это значит (см. с. 77), что вместо «объединенного» электронно-дырочного уравнения (1-102) можно пользоваться уравнением (1-78a), положив $\partial E/\partial x = 0$.

2. Концентрация дырок, инжектируемых в базу, невелика, т. е. выполняется условие низкого уровня инжекции (1-109). При этом полная концентрация дырок в базе ($p = p_0 + \Delta p$) остается значительно меньше концентрации электронов ($n = n_0 + \Delta n$).

Следовательно, согласно (1-76) можно пренебречь дрейфовой составляющей дырочного тока в базе. По аналогичным причинам можно пренебречь дрейфовой составляющей электронного тока в эмиттере. Соответственно вместо уравнения непрерывности (1-78а) можно пользоваться уравнением диффузии (1-79а).

3. Падение напряжения в нейтральном слое базы (а тем более эмиттера) значительно меньше внешнего напряжения, так что последнее можно считать приложенным непосредственно к переходу. Это условие выполняется при достаточно малых токах и сопротивлении базы.

4. Ширина перехода настолько мала, что процессами генерации и рекомбинации в области перехода можно пренебречь. Это дает право считать электронные токи на обеих границах перехода одинаковыми; то же самое относится к дырочным токам.

5. Обратные напряжения значительно меньше напряжения пробоя, так что можно пренебречь предпробойными явлениями в переходе.

6. Отсутствуют всякого рода поверхностные утечки, шунтирующие переход, а следовательно, и токи утечки, которые, вообще говоря, добавляются к токам инжекции и экстракции.

Учитывая принятые допущения, работу диода можно описать следующим образом. При прямом смещении перехода концентрация дырок на его базовой границе повышается, и эти избыточные дырки

диффундируют в глубь базы. По мере удаления от перехода концентрация дырок убывает и в установившемся режиме получается некоторое *распределение* избыточных дырок $\Delta p(x)$ (рис. 2-19, а). Инжекция дырок в базу нарушает ее нейтральность и вызывает приток избыточных электронов из внешней цепи. Эти электроны распределяются таким образом, чтобы компенсировать поле дырок, т. е. накапливаются в той же области, что и дырки. Поэтому кривые $\Delta p(x)$ и $\Delta n(x)$ оказываются почти одинаковыми¹. Небольшая разница между этими кривыми обусловлена различием подвижно-

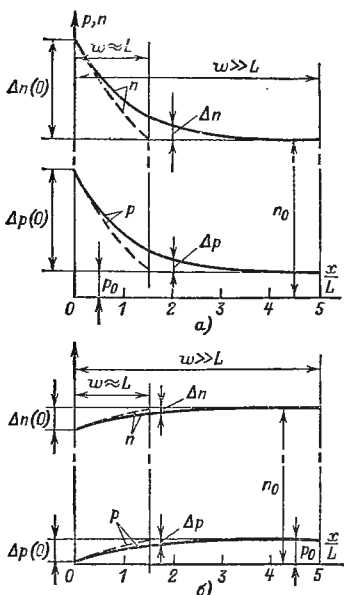


Рис. 2-19. Распределение носителей в диодах с толстой и тонкой базами.

а — при прямом смещении; б — при обратном смещении.

¹ Поскольку в n -базе $n \gg p$, кривые $n(x)$ должны были бы лежать намного выше, чем кривые $p(x)$. Чтобы избежать этого, на рис. 2-19 сделан разрыв на оси ординат.

стей электронов и дырок (эффект Дембера, см. с. 76). В установившемся режиме в базе протекает диффузионный дырочный ток, который согласно (1-73а) пропорционален градиенту концентрации в каждой точке кривой $\Delta p(x)$. Очевидно, что этот ток уменьшается с удалением от перехода. Так как полный ток диода должен быть одинаковым в любом сечении, то уменьшение диффузионного дырочного тока сопровождается ростом электронной составляющей. Структура полного тока рассмотрена в § 2-8. Однако значение полного тока удастся вычислить без учета этой структуры, если воспользоваться сделанными выше допущениями.

Действительно, учитывая допущения 2 и 4, можем записать для базовой границы перехода:

$$j(0) = j_p(0) + j_n(0) = j_{p \text{ диф}}(0) + j_{n \text{ дыф}}(-l),$$

где координата $x = -l$ соответствует эмиттерной границе. В одномерном случае (см. рис. 2-8) плотность тока $j(0)$ сохраняется в любом сечении; в неоднородном случае (когда площадь сечения есть функция координаты x) плотность тока непостоянна, но полный ток $I = j(0) S(0)$ сохраняет свое значение в любом сечении. Реальные структуры полупроводниковых диодов и транзисторов неоднородные; однако анализ (во избежание серьезных математических трудностей), как правило, проводится применительно к одномерной модели (в данном случае применительно к рис. 2-8), после чего в случае необходимости делаются те или иные поправки на неоднородность.

Таким образом, чтобы рассчитать ток диода, нужно, зная приложенное напряжение, найти распределения дырок в базе и электронов в эмиттере, определить градиенты этих распределений соответственно в точках $x = 0$ и $x = -l$ и затем с помощью формул (1-73) получить компоненты полного тока $j_{p \text{ диф}}(0)$ и $j_{n \text{ дыф}}(-l)$. Эта задача решается в следующем разделе применительно к дырочному компоненту — главному в несимметричном переходе p^+-n .

Решение диффузионного уравнения. Чтобы получить статическую вольт-амперную характеристику диода, нужно найти стационарное распределение дырок в базе. Для этого в уравнении диффузии (1-79а) следует положить $\partial \Delta p / \partial t = 0$, после чего оно легко приводится к виду (1-110б):

$$\frac{d^2(\Delta p)}{dx^2} - \frac{\Delta p}{L^2} = 0. \quad (2-26)$$

Здесь для простоты опущен индекс p при параметре L [см. (1-111)].

Как известно, решением (2-26) является сумма двух экспонент:

$$\Delta p(x) = A_1 e^{\frac{x}{L}} + A_2 e^{-\frac{x}{L}}. \quad (2-27)$$

Для того чтобы в решении (2-27) определить коэффициенты A_1 и A_2 , нужно знать граничные условия. В § 1-13 (разд. «Моно-

полярная диффузия») значение $\Delta p(0)$ было бы бесконечно большим и соответственно полагалось $\Delta p(\infty) = 0$. Теперь, учитывая допущение п. 3, выразим граничную концентрацию $\Delta p(0)$ через приложенное напряжение с помощью (2-14а):

$$\Delta p(0) = p_0 \left(e^{\frac{U}{\Phi T}} - 1 \right). \quad (2-28a)$$

Вторую граничную концентрацию запишем в виде

$$\Delta p(w) = (0), \quad (2-28б)$$

считая, что концентрации носителей на омическом контакте (см. с. 112) сохраняют равновесное значение. При граничных условиях (2-28) коэффициенты A_1 и A_2 имеют значения:

$$\left. \begin{aligned} A_1 &= \frac{\Delta p(0)}{2 \operatorname{sh} \left(\frac{w}{L} \right)} e^{-\frac{w}{L}}; \\ A_2 &= \frac{\Delta p(0)}{2 \operatorname{sh} \left(\frac{w}{L} \right)} e^{\frac{w}{L}}, \end{aligned} \right\} \quad (2-29)$$

а распределение $\Delta p(x)$ принимает вид:

$$\Delta p(x) = p_0 \left(e^{\frac{U}{\Phi T}} - 1 \right) \frac{\operatorname{sh} \left(\frac{w-x}{L} \right)}{\operatorname{sh} \left(\frac{w}{L} \right)}. \quad (2-30)$$

Для толстой базы [$w > (2-3) L$] можно положить $w \rightarrow \infty$. Тогда коэффициенты A_1 и A_2 упрощаются:

$$A_1 = 0; \quad A_2 = \Delta p(0),$$

а распределение дырок оказывается экспоненциальным [см. (1-113)]:

$$\Delta p(x) = p_0 \left(e^{\frac{U}{\Phi T}} - 1 \right) e^{-\frac{x}{L}}. \quad (2-31a)$$

В случае тонкой базы ($w < 0,5 L$) можно положить $\operatorname{sh} z \approx z$. Тогда из (2-30) получаем почти линейное распределение:

$$\Delta p(x) = p_0 \left(e^{\frac{U}{\Phi T}} - 1 \right) \left(1 - \frac{x}{w} \right), \quad (2-31б)$$

которое характерно для реальных диодов.

На рис. 2-19, а распределения (2-31а) и (2-31б) показаны соответственно сплошными и пунктирными линиями.

Для режима экстракции ($U < 0$) все выведенные формулы остаются в силе, а соответствующие распределения дырок и электронов показаны на рис. 2-19, б.

Вольт-амперная характеристика. Дифференцируя (2-30) по координате x и подставляя результат в (1-73а), получаем распределение плотности дырочного тока в базе:

$$j_p(x) = \frac{qD_{p6}p_{06}}{L_{p6}} \left(e^{\frac{U}{\Phi_T}} - 1 \right) \frac{\operatorname{ch} \frac{w_6 - x}{L_{p6}}}{\operatorname{sh} \frac{w_6}{L_{p6}}}. \quad (2-32a)$$

Здесь для определенности введены индексы n для базового слоя и p для «рабочих» носителей — дырок. По аналогии можно записать для плотности электронного тока в эмиттерном слое:

$$j_n(x) = \frac{qD_{n3}n_{03}}{L_{n3}} \left(e^{\frac{U}{\Phi_T}} - 1 \right) \frac{\operatorname{ch} \frac{w_3 - x}{L_{n3}}}{\operatorname{sh} \frac{w_3}{L_{n3}}}, \quad (2-32б)$$

где L_{n3} — диффузионная длина электронов в эмиттере, а координата x отсчитывается от перехода в глубь эмиттера.

Полагая $x = 0$ в формулах (2-32), умножая обе части на площадь S и складывая токи $I_p(0)$ и $I_n(0)$, получаем искомую вольт-амперную характеристику идеализированного диода. Обычно ее записывают в следующей форме:

$$I = I_0 \left(e^{\frac{U}{\Phi_T}} - 1 \right), \quad (2-33)$$

где

$$I_0 = \frac{qD_{p6}S}{L_{p6} \operatorname{th} \left(\frac{w_6}{L_{p6}} \right)} p_{06} + \frac{qD_{n3}S}{L_{n3} \operatorname{th} \left(\frac{w_3}{L_{n3}} \right)} n_{03}. \quad (2-34)$$

Формула (2-33) — одна из важнейших в транзисторной технике — представлена на рис. 2-20 в относительных единицах.

Ток I_0 , определяющий «масштаб» характеристики, называется *тепловым током*. Термин «тепловой» отражает сильную температурную зависимость тока I_0 , а также тот факт, что он равен нулю при абсолютном нуле температуры. Другим распространенным термином является «обратный ток насыщения», происхождение которого связано с тем, что при отрицательном напряжении $|U| \gg \Phi_T$ обратный ток идеализированного диода равен $-I_0$ и не зависит от напряжения.

Параметром, характеризующим относительную роль главной составляющей тока в диоде, является коэффициент инжекции γ . В случае p^+-n перехода этот коэффициент записывается следую-

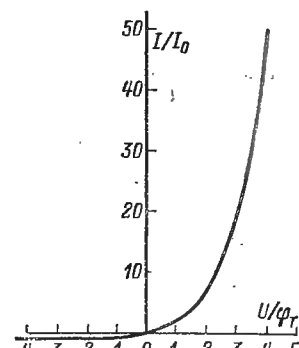


Рис. 2-20. Статическая вольт-амперная характеристика идеализированного плоскостного диода.

щим образом:

$$\gamma = \frac{I_p(0)}{I_p(0) + I_n(0)} = \frac{I_p(0)}{I(0)}.$$

Используя формулы (2-32) при $x = 0$, переходя от коэффициентов диффузии к подвижностям, от концентраций неосновных носителей к концентрациям основных, а затем к удельным сопротивлениям, получаем коэффициент инжекции в следующем виде¹:

$$\gamma = \left[1 + \frac{(\mu_n \mu_p)_a L_{p6} \rho_a}{(\mu_n \mu_p)_b L_{n3} \rho_b} \right]^{-1} \approx 1 - \frac{\rho_a}{\rho_b}. \quad (2-35)$$

Здесь приближение (весьма грубое, но зато и весьма наглядное) основано на том, что коэффициент при ρ_a / ρ_b отличается от единицы не более чем в 2—3 раза, тогда как само отношение удельных сопротивлений редко превышает 0,01. Для значений $\rho_b = 5 \text{ Ом} \cdot \text{см}$ и $\rho_a = 0,01 \text{ Ом} \cdot \text{см}$ получаем $\gamma = 0,998$. Как видим, дырочная составляющая тока в одностороннем p^+n переходе, действительно, является основной. Поэтому вторым слагаемым в (2-34) обычно пренебрегают. Тогда тепловой ток

$$I_0 \approx q \frac{DS}{L \operatorname{th}\left(\frac{w}{L}\right)} p_0. \quad (2-36a)$$

В частности, при $w \ll L$ [когда $\operatorname{th}(w/L) \approx w/L$]

$$I_0 \approx q \frac{DS}{w} p_0. \quad (2-36б)$$

При $w \gg L$ [когда $\operatorname{th}(w/L) \approx 1$]

$$I_0 \approx q \frac{DS}{L} p_0. \quad (2-36в)$$

Свойства теплового тока будут подробно рассмотрены в § 2-6. Здесь только отметим, что так как концентрация p_0 пропорциональна n_i^2 [см. (1-21б)], а собственная концентрация n_i у кремния гораздо меньше, чем у германия, то и тепловой ток у кремниевых диодов несравненно меньше, чем у германиевых.

Одной из важных особенностей характеристики (2-33) является очень крутая (экспоненциальная) прямая ветвь. Поэтому весьма большие прямые токи (несколько ампер и выше) получаются у полупроводниковых диодов при напряжении не более 1 В, т. е. намного

¹ При записи (2-35) приняты соотношения $w_b > L_{p6}$ и $w_a > L_{n3}$, при которых $\operatorname{th}(w/L) \approx 1$. Если для одного из слоев соотношение w и L имеет обратный характер, соответствующая диффузионная длина в (2-35) заменяется на толщину слоя, поскольку в этом случае $\operatorname{th}(w/L) \approx w/L$. В реальных переходах $w_a < L_{n3}$, поэтому в формуле (2-35) вместо отношения L_{p6}/L_{n3} обычно пишут L_{p6}/w_a .

меньшем, чем в случае вакуумных и газонаполненных диодов. В связи с большой крутизной прямой ветви обычно удобнее задавать в качестве аргумента ток, а напряжение считать его функцией. Соответственно формулу (2-33) целесообразно преобразовать к следующему виду:

$$U = \varphi_T \ln \left(\frac{I}{I_0} + 1 \right). \quad (2-37)$$

Вентильные свойства любого диода выражены тем ярче, чем меньше обратный ток при заданном обратном напряжении и чем меньше прямое напряжение при заданном прямом токе. К сожалению, эти два требования в данном случае противоречивы. В самом деле, из формулы (2-37) видно, что изменение теплового тока, какими бы причинами оно не вызывалось, сопровождается изменением прямого напряжения в противоположном направлении.

Это хорошо видно из рис. 2-21, а, где различие токов I_0 обусловлено разницей в площадях переходов при прочих равных условиях. Важным следствием этой общей зависимости является то, что прямые напряжения у кремниевых диодов заметно больше, чем у германиевых, поскольку тепловой ток у первых на несколько порядков меньше. Различие в прямых напряжениях германиевых и кремниевых диодов составляет обычно 0,4 В (рис. 2-21, б) и сохраняется вплоть до таких малых токов [порядка нескольких $(I_0)_{\text{Ge}}$], при которых у германиевых диодов напряжение уже практически равно нулю.

В результате в этом диапазоне токов кремниевым диодам свойствен кажущийся с д и г характеристики по оси напряжений на 0,4 В — так называемая „пятка” (см. пунктирные кривые на рис. 2-21, б, для которых масштаб по оси ординат принят в 1000 раз меньшим, чем для сплошных).

Характеристические сопротивления. Нелинейность характеристики диода удобно оценивать, сопоставляя его сопротивления в прямом и обратном направлениях. Как и для других нелинейных элементов, различают дифференциальные сопротивления и сопротивления постоянному току.

Найдя производную от функции (2-37), легко представить дифференциальное сопротивление диода как функцию тока (рис. 2-22, а):

$$r_d = \frac{dU}{dI} = \frac{\varphi_T}{I + I_0} \approx \frac{\varphi_T}{I}. \quad (2-38)$$

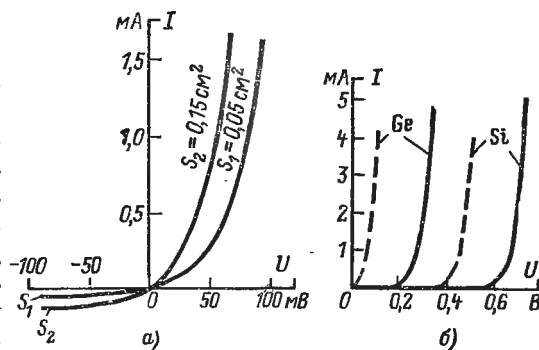


Рис. 2-21. Характеристики идеализированных диодов с разными площадями переходов (а) и разными тепловыми токами — германиевого и кремниевого (б).

Приближенное выражение, размещается, действительно только для прямой ветви при условии $I \gg I_0$. На обратной ветви сопротивление r_d резко возрастает и при $|U| \gg \varphi_T$ может считаться бесконечно большим. На прямой ветви сопротивление r_d , наоборот, быстро уменьшается и, например, при токе $I > 5 \div 10$ мА составляет несколько ом. При таком сопротивлении изменение прямого напряжения даже на 5—10 мВ приводит к значительным изменениям тока. Поэтому задать прямое напряжение с целью получить нужный ток весьма трудно и для полупроводникового диода, работающего в прямом направлении, более характерен режим заданного тока. На это указывалось при выводе формулы (2-37).

Часто при расчете диодных схем пользуются сопротивлениями постоянному току. Из формулы (2-37) получаем зависимость такого сопротивления от тока (рис. 2-22, а):

$$R_d = \frac{U}{I} = \frac{\varphi_T}{I} \ln \left(\frac{I}{I_0} + 1 \right). \quad (2-39a)$$

Из формулы (2-33) получаем зависимость этого же сопротивления от напряжения (рис. 2-22, б):

$$R_d = \frac{U}{I_0 (e^{U/\varphi_T} - 1)}. \quad (2-39б)$$

Отсюда следует, что на обратной ветви характеристики, когда $|U| \gg \varphi_T$, сопротивление R_d прямо пропорционально напряжению:

$$R_d = \frac{|U|}{I_0}.$$

В нулевой точке ($I = 0$; $U = 0$) сопротивления r_d и R_d , как нетрудно убедиться, имеют одно и то же значение:

$$r_{d0} = R_{d0} = \frac{\varphi_T}{I_0}. \quad (2-40)$$

Например, при токе $I_0 = 1$ мкА $r_{d0} = R_{d0} = 25$ кОм. У кремниевых диодов, имеющих гораздо меньший ток I_0 , сопротивление $r_{d0} = R_{d0}$ должно было бы составлять сотни мегаом. На самом деле оно меньше по причинам, изложенным в § 2-6.

В области прямых токов сопротивление R_d всегда больше сопротивления r_d , а в области обратных токов оно всегда меньше сопротивления r_d (рис. 2-22). Обычно (если не делают специальных оговорок) под прямым сопротивлением диода $R_{пр}$ понимают сопротивление R_d , соответствующее номинальному прямому току (или напряжению), указанному в справочнике. Под обратным сопротивлением диода $R_{обр}$ обычно подразумевают сопротивление R_d при номинальном обратном напряжении диода.

Необходимо подчеркнуть, что формулы (2-38)—(2-40) выведены на основе характеристики (2-33), которая не отражает всех свойств реального диода. Поэтому на практике значения сопротив-

лений r_d и R_d отличаются от значений, вычисленных по указанным формулам. Причины этих различий рассмотрены в § 2-6 и 2-8.

Температура перехода. Поскольку на переходе падает основная часть приложенного напряжения и выделяется основная мощность, температура перехода может заметно отличаться от температуры внешних электродов, корпуса диода и окружающей среды.

Тепловой баланс в системе «диод—внешняя среда» устанавливается через некоторое время после включения диода и определяется теплопроводностью отдельных элементов системы. Для практических расчетов важно знать связь между температурой перехода $T_{пер}$, которую трудно измерить непосредственно, и температурой окружающей среды $T_{окр}$, которая обычно известна. Такая связь дается простым выражением

$$T_{пер} - T_{окр} = R_t P, \quad (2-41)$$

где R_t — так называемое *тепловое сопротивление* ($^{\circ}\text{C}/\text{Вт}$), а $P = UI$ — мощность, выделяемая в переходе

Тепловое сопротивление, подобно электрическому, выражается формулой

$$R_t = \frac{1}{\lambda} \frac{l}{S},$$

где λ — удельная теплопроводность (играющая роль удельной проводимости), а l и S — длина и площадь того слоя, через который протекает тепловой поток. Количество таких слоев в реальных конструкциях полупроводниковых приборов достаточно велико, и они имеют весьма различные конфигурации и удельные теплопроводности. Например, в диоде, помимо p - и n -слоев (с резко различными толщиной и площадью), имеются еще слой клея или припоя, крепящего кристалл к «ножке», сама «ножка», металлические контакты и т. д. Поэтому результирующее тепловое сопротивление диода является суммой сопротивлений отдельных слоев. Сколько-нибудь точный расчет теплового сопротивления невозможен; его, как правило, оценивают из эксперимента.

Для маломощных приборов (50—100 мВт) с площадью кристалла примерно 1 мм^2 характерны значения $R_t = 600 \div 800^{\circ}\text{C}/\text{Вт}$, для приборов средней мощности (примерно 1 Вт) с соответственно большей площадью кристалла — значения $R_t = 50 \div 150^{\circ}\text{C}/\text{Вт}$, а для мощных приборов (10—20 Вт и более) — значения $R_t = 2 \div 5^{\circ}\text{C}/\text{Вт}$.

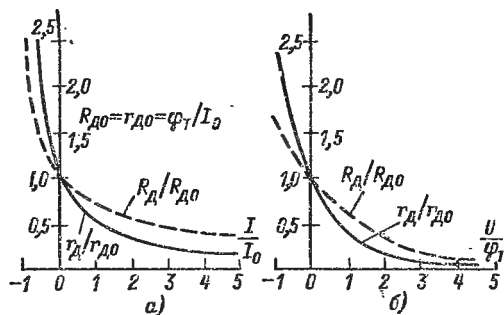


Рис. 2-22. Зависимость сопротивлений идеализированного диода — дифференциального сопротивления r_d и сопротивления постоянному току R_d — от прямого тока (а) и прямого напряжения (б).

Обычно полное тепловое сопротивление прибора разбивается на два слагаемых: тепловое сопротивление участка «переход—корпус» $R_{тк}$ и тепловое сопротивление участка «корпус—среда» $R_{тс}$. Последнее в решающей степени определяется способом теплоотвода и охлаждения (см. § 4-11, рис. 4-34).

В момент подачи прямого тока температура p - n перехода равна температуре окружающей среды и лишь постепенно повышается до установившегося значения, определяемого формулой (2-41). Такая тепловая инерционность прибора обычно характеризуется постоянной времени нагревания или охлаждения τ_t . Сам переходный процесс изменения температуры считается экспоненциальным:

$$\Delta T(t) = \Delta T(\infty) (1 - e^{-t/\tau_t}).$$

Тепловая постоянная времени, как и тепловое сопротивление, является интегральным параметром: она складывается из постоянных времени отдельных слоев прибора. Для каждого из слоев

$$\tau_t = CR_t,$$

где C — теплоемкость слоя. Выражение для теплоемкости

$$C = c\rho V$$

показывает, что она, помимо удельной теплоемкости c и удельного веса ρ , пропорциональна объему слоя V . Следовательно, можно наперед предсказать, что результирующая постоянная времени будет определяться наиболее массивным конструктивным элементом прибора — корпусом или (если он есть) охлаждающим радиатором (§ 4-11). Например, непосредственно для области p - n перехода постоянная времени для диодов средней мощности составляет 15—25 мс, а в целом (для диода в корпусе) — 4—8 с. При наличии радиатора постоянная времени возрастает до 100—200 с и более.

Заметим, что постоянная времени p - n перехода соизмерима с периодом сетевой частоты (20 мс). Это значит, что при сетевой частоте температура перехода «не успевает следить» за изменениями мгновенной мощности в нем. Последствия такой инерционности иногда существенны (см. § 12-3).

Более подробно вопросы теплового режима рассмотрены в [30].

2-6. ОБРАТНАЯ ХАРАКТЕРИСТИКА РЕАЛЬНОГО ДИОДА

Опыт показывает, что обратный ток диода не остается постоянным и равным I_0 , как следует из формулы (2-33). Он всегда более или менее значительно растет с увеличением обратного напряжения. Но и при малых напряжениях обратный ток больше теплового, особенно у кремниевых диодов, у которых это превышение достигает 2—3 порядков. Главные причины отклонения реальной обратной характеристики от идеализированной заключаются в термогенерации носителей в области перехода, поверхностных утечках, а также в явлениях, которые при достаточно большом обратном

напряжении приводят к пробое переходов. Таким образом, обратный ток диода состоит из нескольких компонентов, из которых при выводе характеристики (2-33) мы познакомились лишь с одним — тепловым током. Однако и этот ток следует рассмотреть подробнее.

Тепловой ток. Используя в формуле (2-36в) соотношения (1-16), (1-36) и (1-74), нетрудно представить тепловой ток в следующем виде:

$$I_0 = \frac{b}{(b+1)^2} \frac{\varphi_T S \rho_0}{L \rho_i^{\frac{3}{2}}},$$

где $b = \mu_n / \mu_p$. Например, для германиевых и кремниевых диодов при $\rho_0 = 5 \text{ Ом} \cdot \text{см}$, $\tau = 5 \text{ мкс}$ и $S = 0,01 \text{ см}^2$ получим $(I_0)_{\text{Ge}} = 6 \text{ мкА}$; $(I_0)_{\text{Si}} = 0,6 \cdot 10^{-6} \text{ мкА}$, т. е. различие в токах I_0 составляет около семи порядков. Малый тепловой ток кремниевых диодов является одним из их важнейших преимуществ, хотя полный обратный ток реальных кремниевых диодов, как будет видно ниже, значительно больше, чем I_0 .

Запишем формулу (2-34) в несколько иной форме. Умножим и разделим слагаемые в правой части на соответствующие диффузионные длины, положим $\text{th}(\omega/L) = 1$ и используем соотношение (1-111). Тогда

$$I_0 = q (SL_{p0}) \frac{\rho_{00}}{\tau_0} + q (SL_{n0}) \frac{n_{00}}{\tau_0}. \quad (2-42)$$

Поскольку множители ρ_{00}/τ_0 и n_{00}/τ_0 суть скорости генерации дырок в базе и электронов в эмиттере [см. (1-43)], выражение (2-42) можно трактовать следующим образом: *тепловой ток обусловлен генерацией неосновных носителей в объемах SL , прилегающих к переходу*, откуда эти носители диффундируют в область потенциального барьера и уносятся полем в другой слой (рис. 2-23). В равновесном состоянии эти потоки компенсируются встречными потоками аналогичных носителей, имеющих достаточную энергию, чтобы преодолеть потенциальный барьер.

Особый интерес представляет температурная зависимость теплового тока. Если пренебречь электронной составляющей, то согласно выражению (2-36в) тепловой ток зависит от температуры через параметры D , τ , ρ_0 . Из них главную роль играет равновесная концентрация неосновных носителей ρ_0 . Принимая для нее выражение (1-21б), получаем соотношение $I_0 \sim n_i^2$; подставляя значение n_i из (1-15), выразим тепловой ток в следующем виде:

$$I_0(T) = I_{00} e^{-\varphi_0/\varphi_T}. \quad (2-43a)$$

Здесь ток I_{00} содержит величины, мало зависящие от температуры. С ростом температуры примесный полупроводник постепенно превращается в собственный (с. 33). Следовательно, при темпера-

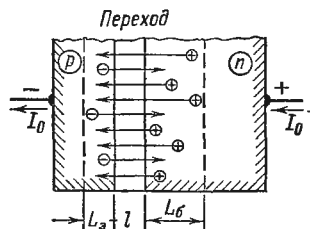


Рис. 2-23. Происхождение электронного и дырочного компонентов теплового тока.

турах выше критической [см. (1-226)] можно считать $p_0 \approx n_i$ и соответственно $I_0 \sim n_i$. В этом случае с учетом (1-15) выражение для теплового тока примет вид:

$$I_0(T) = I_{00} e^{-\varphi_3/2\varphi T}. \quad (2-43б)$$

На практике всегда известен тепловой ток при некоторой (обычно «комнатной») температуре T_0 и требуется определить его значение при другой температуре T . Из формулы (2-43а) легко получить:

$$\frac{I_0(T)}{I_0(T_0)} = e^{\varphi_3 \left(\frac{1}{\varphi T_0} - \frac{1}{\varphi T} \right)}. \quad (2-44)$$

Разность, стоящую в круглых скобках показателя степени, можно привести к общему знаменателю и подставить значения φT_0 и φT из (1-3); тогда

$$\frac{1}{\varphi T_0} - \frac{1}{\varphi T} = 11\,600 \frac{T - T_0}{TT_0} = 11\,600 \frac{\Delta T}{TT_0}.$$

Используя это преобразование и полагая, что абсолютная температура в рабочем диапазоне меняется не очень сильно (т. е. $TT_0 \approx T_0^2$), получаем вместо (2-44) простое приближенное соотношение:

$$I_0(T) \approx I_0(T_0) e^{a \Delta T}, \quad (2-45)$$

где $a = 0,13 \varphi_3$:

$$a_{Si} = 0,13^\circ\text{C}^{-1}; \quad a_{Ge} = 0,09^\circ\text{C}^{-1}.$$

Этими значениями коэффициента a можно пользоваться вплоть до температур $120\text{—}150^\circ\text{C}$ для кремния и $70\text{—}80^\circ\text{C}$ для германия. При более широком температурном диапазоне пользуются усредненными значениями $a_{Si} \approx 0,1$ и $a_{Ge} \approx 0,07$.

Несмотря на простоту формула (2-45) не всегда удобна для быстрых «прикидок». Поэтому целесообразно заменить степень числа e степенью числа 2, которую всегда легко вычислить в уме. Заменив основание степени по известным правилам, формулу (2-45) можно представить в следующем виде:

$$I_0(T) \approx I_0(T_0) \cdot 2^{\Delta T/T^*}, \quad (2-46)$$

где параметр $T^* = (\ln 2)/a$ можно назвать *температурой удвоения* тока; это приращение температуры, при котором тепловой ток удваивается. Например, при $a = 0,07; 0,09; 0,1; 0,13$ получаем соответственно $T^* = 10, 8, 7, 5^\circ\text{C}$. На практике распространено правило: «тепловой ток удваивается на каждые 10°C приращения температуры», что соответствует значению $a = 0,07$. Ясно, что это правило не универсально и обычно занижает фактические изменения теплового тока в несколько раз.

Ток термогенерации. В идеализированном диоде мы считали переход бесконечно узким и, следовательно, могли пренебречь генерацией и рекомбинацией носителей в этой области (см. п. 4

на с. 114). Реальный переход имеет конечную ширину, поэтому указанные процессы имеют в нем место так же, как в любом другом слое полупроводника, и играют немаловажную роль [31].

Электрическое поле, которое всегда есть в переходе, быстро уносит генерируемые носители в соответствующий слой диода, что вызывает протекание некоторого тока — *тока термогенерации* I_G . В равновесном состоянии диода этот ток компенсируется равным ему встречным током — *током рекомбинации* I_R . Ток рекомбинации обусловлен теми носителями, которые непрерывно проникают в переход из эмиттера и базы, но не имеют достаточной энергии, чтобы перейти в смежный слой. Вблизи точки «отражения» (см. рис. 2-6) такие носители имеют малую скорость и успевают рекомбинировать.

В неравновесном состоянии диода взаимная компенсация токов I_G и I_R нарушается. В случае обратного включения диода превалирует ток термогенерации, так как высота потенциального барьера увеличивается и проникание носителей в область перехода (с последующей рекомбинацией) затрудняется. При обратном смещении $|U| \gg \phi_T$ практически остается только ток термогенерации, пропорциональный объему «генерирующего» обедненного слоя, т. е. ширине перехода (рис. 2-24). Ток I_G накладывается на тепловой ток I_0 , и результирующий обратный ток оказывается больше, чем это следует из формулы (2-34). Более того, ток I_G растет с расширением перехода, т. е. с ростом напряжения, что приводит к конечному наклону обратной ветви характеристики.

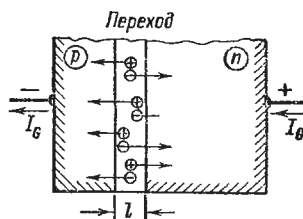


Рис. 2-24. Происхождение тока термогенерации в переходе.

Для того чтобы оценить значение тока термогенерации, воспользуемся той трактовкой, которая была дана для теплового тока в связи с формулой (2-42). В области перехода при его обратном смещении концентрации n и p близки к нулю. Если принять $n_i = p_i = n_i$, т. е. положить уровни ловушек расположенными в середине запрещенной зоны, то скорость генерации будет равна n_i/τ_{∞} , где $\tau_{\infty} = \tau_{p0} + \tau_{n0}$ [см. (1-60в)]. Соответственно ток термогенерации по аналогии с выражением (2-42) запишется в следующем виде:

$$I_G = q(Sl) \frac{n_i}{\tau_{\infty}}, \quad (2-47)$$

где l — ширина перехода.

Пусть, например, $S = 0,01 \text{ см}^2$; $l = 1,25 \text{ мкм}$; $\tau_{\infty} = 4 \text{ мкс}$; тогда из формулы (2-47) для кремниевых диодов получаем $I_G = 0,001 \text{ мкА}$. Для германиевых диодов ток I_G будет в 1000 раз больше, т. е. около 1 мкА .

Сравним токи I_G и I_0 . Разделив (2-47) на первый член (2-42), выразив концентрации через удельные сопротивления и полагая для простоты $\tau_\infty = 2\tau_6$ и $\mu_n = \mu_p$, получим:

$$\frac{I_G}{I_0} \approx \frac{\rho_l}{\rho_6} \frac{l}{L}. \quad (2-48)$$

Для германия при $\rho_6 = 5 \text{ Ом} \cdot \text{см}$, $l = 1 \text{ мкм}$, $L = 150 \text{ мкм}$ и комнатной температуре отношение токов составит около 0,1. Для кремния при прочих равных условиях отношение токов составит около 1000. Таким образом, при комнатной температуре ток термогенерации в германиевых диодах пренебрежимо мал и их обратный ток близок к тепловому; в кремниевых диодах, наоборот, ток термогенерации является главным компонентом обратного тока. В связи с этим обстоятельством различие в полных обратных токах у германиевых и кремниевых диодов получается не столь большим, как следует из формулы (2-34). Разница в несколько тысяч раз (при малом напряжении) — типичная величина.

Что касается зависимости тока термогенерации от напряжения, то она получается при подстановке ширины перехода l из формулы (2-12) в выражение (2-47) и имеет вид:

$$I_G \sim \sqrt{|U|}. \quad (2-49)$$

При этом легко убедиться, что сопротивления обратной ветви r_d и R_d будут тоже пропорциональны величине $\sqrt{|U|}$.

Ток термогенерации I_G , как видно из выражения (2-47), пропорционален собственной концентрации в первой степени и, следовательно, с точки зрения температурной зависимости, описывается формулой типа (2-43б). Соответственно коэффициенты a [см. формулу (2-45)] равны:

$$a_{Si} \approx 0,07; \quad a_{Ge} \approx 0,05,$$

а температуры удвоения тока [см. формулу (2-46)]

$$T_{Si}^* \approx 10^\circ\text{C}; \quad T_{Ge}^* \approx 14^\circ\text{C}.$$

Заметим, что у кремниевых диодов ток термогенерации является главным компонентом обратного тока при комнатной температуре. С повышением температуры тепловой ток I_0 растет быстрее (так как для него $T_{Si}^* \approx 5^\circ\text{C}$) и в конце концов начинает превышать ток I_G . Обычно это происходит при температуре $+100^\circ\text{C}$ и выше. У германиевых диодов при комнатной температуре доминирует тепловой ток, а ток термогенерации начинает играть роль лишь при отрицательной температуре. Однако в этом диапазоне значение обратного тока делается вообще малосущественным.

Поверхностные каналы. В § 2-4 отмечалось, что наличие поверхностных энергетических уровней приводит к обогащению или обеднению приповерхностного объема основными носителями. Нередко, особенно в высокоомных полупроводниках, наблюдается случай, когда под действием поверхностных уровней образуется не только

обедненный слой, но и тонкий приповерхностный канал — и н в е р с и о н н ы й слой (см. с. 71 и 110). Если такое явление имеет место в базе диода, то структура p - n перехода существенно изменится (ср. рис. 2-25, а и б). А именно, переход оказывается расположенным не только вблизи металлургической границы, но и под поверхностью вдоль базы. Значит, при наличии канала площадь перехода увеличивается, а вместе с нею, естественно, возрастает обратный ток.

Поскольку часть обратного тока протекает по каналу, потенциал последнего повышается по мере удаления от металлургической границы. Соответственно разность потенциалов между объемом базы и каналом уменьшается, а ширина приповерхностного перехода

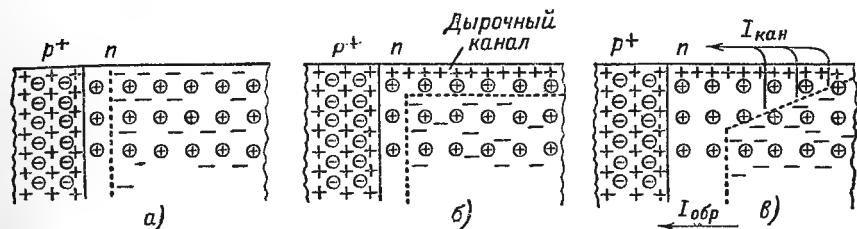


Рис. 2-25. Структура p - n перехода в отсутствие (а) и при наличии (б и в) поверхностного канала. Базовая граница перехода показана точками.

убывает (рис. 2-25, в). В той точке, где разность потенциалов падает до величины $\Delta\phi_0$, приповерхностный переход находится в квазиравновесном состоянии и обратный ток делается равным нулю. Следовательно, «рабочая» часть канала, обуславливающая увеличение обратного тока, имеет к о н е ч н у ю длину (обычно десятки доли миллиметра), т. е. меньше физической длины канала.

С ростом обратного напряжения растет «рабочая» длина канала, а значит, и обратный ток приповерхностного перехода. Анализ показывает [15], что функция $I_{\text{кан}}(U_{\text{обр}})$ логарифмическая, т. е. не очень сильная. Однако она приводит к к о н е ч н о м у на к л о н у обратной характеристики независимо от наличия или отсутствия тока термогенерации (т. е. и у кремниевых, и у германиевых диодов).

Заметим, что наличие поверхностных уровней не всегда сопровождается образованием канала. При недостаточной плотности этих уровней образуется только обедненный слой, что равносильно простому увеличению площади p - n перехода. Соответственно увеличатся и ток термогенерации, и тепловой ток; если последний превалирует (у германиевых диодов), то зависимость обратного тока от напряжения будет отсутствовать. Наконец, возможен случай, когда плотность поверхностных уровней недостаточна даже для образования обедненного слоя (точнее, для образования области пространственного заряда п р и м е с е й). Тогда вблизи поверхности просто будет повышенное удельное сопротивление базы и, значит, в этой области ширина перехода будет больше, чем вдали от поверхности.

Ток утечки. Поверхностные утечки представляют собой нередко главный фактор, влияющий на обратную характеристику. Ток утечки не всегда является второстепенным результатом «загрязнения» поверхности. Он обусловлен в первую очередь поверхностными энергетическими уровнями, которые способствуют активной генерации — рекомбинации (см. с. 46), а также молекулярными или ионными пленками, шунтирующими переход (это могут быть молекулы окислов основного материала, молекулы газов, воды, ионы водорода и т. п.). При повышении напряжения ток утечки растет сначала почти линейно, а затем более сильно (рис. 2-26). Почти линейный начальный участок характеристики $I_y(U)$ можно охарактеризовать эквивалентным сопротивлением утечки R_y .

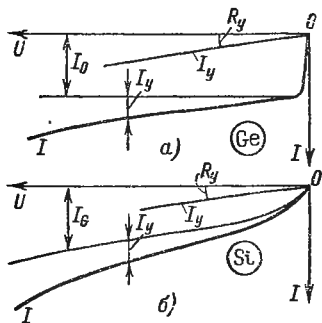


Рис. 2-26. Обратные характеристики диодов — германниевый (а) и кремниевый (б) — в отсутствие поверхностного канала.

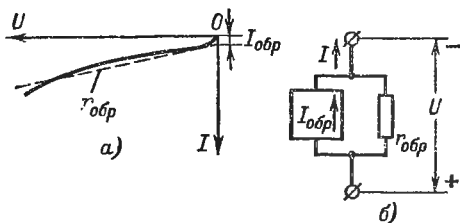


Рис. 2-27. Обратная характеристика реального диода, ее идеализация (а) и эквивалентная схема диода при обратном включении (б).

Характерная черта тока утечки заключается в его временной нестабильности, которую часто называют «ползучестью». Ползучесть проявляется в изменении обратного тока в течение некоторого времени после скачкообразного изменения обратного напряжения, в частности после его включения [32]. Есть основания считать, что ползучесть связана главным образом с адсорбированной пленкой водяных паров. Времена релаксации при нарастании или спадании обратного тока оказываются различными и обычно лежат в пределах от нескольких секунд до нескольких часов. В течение указанного времени обратный ток заметно меняется (изменение может выражаться в десятках процентов), причем величина ползучести оказывается индивидуальной у разных диодов одного и того же типа.

Ток утечки зависит от температуры сравнительно слабо. Поэтому по сравнению с токами I_0 и I_G ток I_y можно считать постоянным. Если при комнатной температуре $I_y \ll I_0 + I_G$, то ролью тока утечки вообще можно пренебречь. Однако чаще, особенно у кремниевых диодов, имеет место соотношение $I_y > I_0 + I_G$.

Эквивалентная схема диода при обратном смещении. Из предыдущих разделов видно, что полный обратный ток диода представляет собой сложную функцию напряжения и возрастает с его ростом (рис. 2-27, а). Для расчетов удобно представить полупроводниковый

вый диод, работающий в обратном направлении, в виде линейной эквивалентной схемы, показанной на рис. 2-27, б Соответствующая формула для такой идеализированной характеристики имеет вид:

$$I = I_{\text{обр}} + \frac{U}{r_{\text{обр}}}, \quad (2-50)$$

где $I_{\text{обр}}$ — ток, получаемый путем экстраполяции характеристики до пересечения с осью токов (рис. 2-27, а); $r_{\text{обр}}$ — сопротивление, характеризующее «средний» наклон кривой.

Несмотря на приближенность формулы (2-50), она, как и эквивалентная схема на рис. 2-27, б, позволяет производить полезные количественные оценки в широком диапазоне напряжений. Параметры эквивалентной схемы определяются по данным справочников или путем измерений.

2-7. ПРОБОЙ ПЕРЕХОДА

Под пробоем p - n перехода понимают резкое уменьшение д и ф е р е н ц и а л ь н о г о обратного сопротивления, сопровождающееся резким возрастанием обратного тока при незначительном увеличении напряжения. Разумеется, понятие «резких» изменений условно; по существу те процессы, которые обуславливают пробой, начинают проявляться в той или иной мере при напряжениях, значительно меньших пробивного. Поэтому, рассматривая причины пробоя, мы тем самым рассмотрим дополнительные причины, по которым обратный ток реального диода превышает тепловой ток I_0 .

Различают три вида (механизма) пробоя: *туннельный* (зенеровский)¹, *лавинный* и *тепловой*. Первые два связаны с увеличением напряженности электрического поля, а третий — с увеличением рассеиваемой мощности и соответственно температуры.

Туннельный пробой. В основе этого вида пробоя лежит туннельный эффект, т. е. «просачивание» электронов сквозь потенциальный барьер, если толщина последнего достаточно мала.

Вероятность туннельного эффекта определяется экспонентой [29]:

$$\exp\left(-\frac{4\pi}{k} d \sqrt{2m^* q \Phi}\right) = \exp(-10^8 d \sqrt{\Phi}),$$

где Φ — высота барьера, В; d — его толщина, см. Коэффициент 10^8 получен для случая $m^* = m$.

При анализе пробоя под высотой барьера следует понимать ширину запрещенной зоны Φ_z , а под толщиной барьера — расстояние d между «противостоящими» зонами (рис. 2-28, а). Если распределение потенциала принять линейным и положить $|U| \gg \Delta\Phi_0$,

¹ Второе название связано с фамилией ученого Zener, впервые описавшего соответствующее явление применительно к однородному материалу.

то состояние d легко найти из рис. 2-28, а: $d = l (\varphi_3 / |U|)$. Тогда вероятность туннелирования (см. выше) будет определяться напряженностью поля в переходе ($E = |U|/l$), шириной запрещенной зоны, а также эффективной массой носителей. Нетрудно убедиться в очень сильной зависимости туннельного эффекта от напряженности поля: изменение значения E от 10^5 до $1,1 \cdot 10^5$ В/см (при $\varphi_3 = 1$ В) увеличивает вероятность туннелирования в e^{100} раз! Ясно, что при напряженностях 10^4 В/см и менее нет никакого смысла учитывать туннельный эффект. Напротив, при напряженностях более 10^5 В/см этот эффект становится весьма существенным.

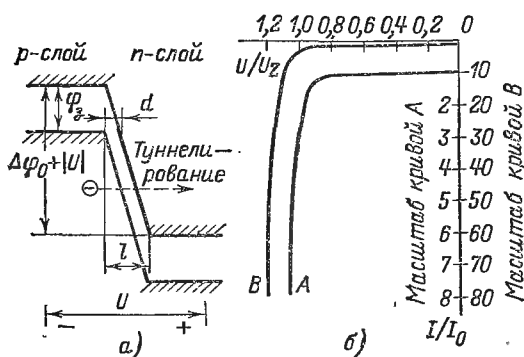


Рис. 2-28. Туннельный пробой.

а — зонная диаграмма; б — обратная характеристика диода в режиме пробоя.

путем умножения приведенной выше вероятности туннелирования на некоторые параметры валентной зоны, а также на напряженность поля. В результате ¹

$$j = AU \exp\left(-\frac{10^8 \varphi_3^{2/2}}{E}\right). \quad (2-51)$$

Начало пробоя оценивают условно, например, из равенства $I = 10I_0$. При этом получают пробивную напряженность поля $E_{пр}$ (опыт дает для германия $E_{пр} \approx 2 \cdot 10^5$ В/см и для кремния $E_{пр} \approx 4 \cdot 10^5$ В/см).

Как известно, в реальных p - n переходах распределение потенциала нелинейное и поле непостоянное: максимальная напряженность поля имеет место на металлургической границе (см. рис. 2-7). Величину $E_{макс}$ можно выразить через напряжение U , если подставить в формулу (2-66) значения $x = 0$ и $l_n = l$, где l определяется формулой (2-12). Полагая $E_{макс} = E_{пр}$, получаем напряжение туннельного пробоя в следующем виде:

$$U_z = \frac{\varepsilon_0 \varepsilon E_{пр}^2}{2qN_6} = \frac{1}{2} \varepsilon_0 \varepsilon \mu_6 E_{пр}^2 \rho_6. \quad (2-52)$$

Более точные полуэмпирические зависимости имеют вид [34]: для германия

$$U_z \approx 100 \rho_n + 50 \rho_p; \quad (2-53a)$$

¹ В несколько иной форме такое выражение получено в работе [33].

$$U_z \approx 40\rho_n + 8\rho_p,$$

$$(2-536)$$

где ρ_n и ρ_p — удельные сопротивления соответствующих слоев (Ом·см).

Важно отметить, что напряжение U_z пропорционально удельному сопротивлению базы. Именно поэтому у высоковольтных диодов делают базу из как можно более высокоомного материала. Кроме того, из формул (2-52) и (2-53) видно, что напряжение туннельного пробоя зависит от типа проводимости базы: для базы типа n оно больше, чем для базы типа p , так как $\mu_n > \mu_p$.

Используя выражение (2-51), можно получить дифференциальное сопротивление диода в области пробоя:

$$r_d = \frac{dU}{dI} \approx \frac{U_z}{I} E_{пр} \cdot 10^{-7}.$$

$$(2-54)$$

Для германия при $U_z = 100$ В и $I = 10$ мА получается $r_d = 200$ Ом.

Как видим, это сопротивление мало и уменьшается с увеличением тока. Общий вид обратной характеристики с учетом туннельного пробоя показан на рис. 2-28, б.

Лавинный пробой. Второй механизм пробоя заключается в лавинном «размножении» носителей в сильном электрическом поле [35]. Этот процесс можно представить себе так же, как ударную ионизацию газа. Электрон и дырка (аналог положительного иона в газе), ускоренные полем на длине свободного пробега, могут разорвать одну из валентных связей атома полупроводника, расположенного в области перехода. В результате рождается новая пара электрон — дырка и процесс может повторяться под действием этих новых носителей (рис. 2-29, а). Тогда суммарный обратный ток через переход окажется больше, чем в отсутствие такой ионизации. При достаточно большой напряженности поля, когда исходная пара носителей в среднем порождает несколько более одной новой пары, ионизация может приобрести лавинный характер подобно самостоятельному разряду в газе. При этом ток будет ограничиваться только внешним сопротивлением.

Ход характеристики в области ионизации вплоть до пробоя описывается с помощью полуэмпирической формулы

$$M = \frac{|I|}{I_0} = \frac{1}{1 - \left(\frac{U}{U_m}\right)^n}, \quad (2-55)$$

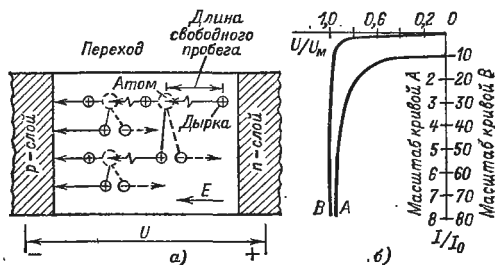


Рис. 2-29. Лавинный пробой.

а — схема размножения дырок; б — обратная характеристика диода в режиме лавинного пробоя.

где M — коэффициент ударной ионизации; U — модуль обратного напряжения и U_M — напряжение лавинного пробоя, при котором $M = \infty$. Значения показателя n приведены в табл. 2-2. График функции $M(U)$ ясен из рис. 2-29, б. Из рисунка видно, что заметный рост обратного тока начинается при $U \approx 0,3 U_M$. Напряжение U_M при лавинном пробое зависит от удельного сопротивления базы; зависимость эта также полуэмпирическая и имеет вид:

$$U_M = a \rho_6^m, \quad (2-56)$$

где ρ_6 имеет размерность Ом·см, а значения a и m приведены в табл. 2-2. Как видим, напряжение лавинного пробоя тоже увеличивается с ростом удельного сопротивления. Сравнивая выражения (2-56) и (2-52), легко прийти к заключению, что отношение $U_Z/U_M = \rho_6^{1-m}$ находится в прямой зависимости от удельного сопротивления базы. При высоких значениях ρ_6 получается $U_Z > U_M$ и пробой носит лавинный характер; при низких значениях ρ_6 получается $U_Z < U_M$ и пробой носит туннельный характер. Граничное значение ρ_6 , при котором $U_Z = U_M$, зависит от материала и типа проводимости. Например, для электронного германия оно составляет около 1 Ом·см.

Таблица 2-2

Параметры, определяющие лавинный пробой

Материал	Тип базы	n	a	m
Кремний	Электронная	5	86	0,65
	Дырочная	3	23	0,75
Германий	Электронная	3	83	0,60
	Дырочная	5	52	0,60

Из выражения (2-55) легко получить дифференциальное сопротивление диода в области лавинного пробоя:

$$r_d = \frac{U}{nI} \frac{1 - (U/U_M)^n}{(U/U_M)^n}. \quad (2-57)$$

Если, например $U_M = 100$ В; $I = 10$ мА; $n = 3$ и $U/U_M = 0,99$, то $r_d = 100$ Ом.

Иногда в литературе можно встретить термин *поверхностный пробой*. Практически этот тип пробоя является либо туннельным, либо лавинным, но происходит в специфических условиях: при сужении p - n перехода вблизи поверхности. Такое сужение может быть обусловлено влиянием поверхностных энергетических уровней. А именно, если наличие последних приводит к обогащению приповерхностной области базы основными носителями¹,

¹ В разделе «Поверхностные каналы» (см. выше) рассмотрен случай обеднения приповерхностной области.

то удельное сопротивление приповерхностного слоя будет меньше, чем в объеме, и ширина перехода в этом слое будет соответственно меньше [см. (2-9)]. Значит, пробой начнется в этом «ослабленном» слое, так как напряжения туннельного и лавинного пробоев согласно (2-52) и (2-56) пропорциональны удельному сопротивлению. Внешние «поверхностный пробой» проявляется в том, что он происходит при напряжении, меньшем «расчетного» (если расчет исходил из объемного удельного сопротивления).

Тепловой пробой. Третий механизм пробоя обусловлен выделением тепла в переходе при протекании обратного тока [36]. Пусть задано обратное напряжение U . Тогда рассеиваемая мощность составит $P = UI_0$. Под действием этой мощности температура перехода повысится на $R_t P$ [(2-41)]. Соответственно возрастут ток I_0

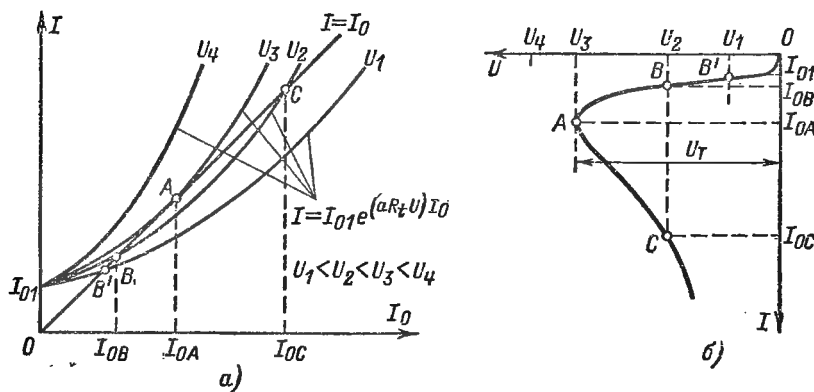


Рис. 2-30. Тепловой пробой.

а — графическое решение уравнения теплового пробоя; *б* — обратная характеристика диода в режиме теплового пробоя.

и мощность P . Такая взаимосвязь может привести к лавинообразному увеличению тока, т. е. к пробое переходу. Оценим условия лавинообразного процесса. Поскольку и тепловой ток, и ток термогенерации описываются выражением (2-45), подставим в последнее $\Delta T = (UI_0) R_t$ и обозначим $I_0(T_0) = I_{01}$. Тогда получаем трансцендентное уравнение относительно тока I_0 :

$$I_0 = I_{01} e^{(a U R_t) I_0}. \quad (2-58)$$

Из рис. 2-30, *а* видно, что это уравнение в зависимости от значения U может иметь либо два корня, либо один кратный корень. При достаточно большом U уравнение вообще не имеет корней. Такие свойства соответствуют обратной характеристике, показанной на рис. 2-30, *б* и имеющей участок с отрицательным дифференциальным сопротивлением. Очевидно, что напряжению теплового пробоя U_T соответствует на рис. 2-30, *а* точка касания *А*, в которой производные обеих частей

уравнения (2-58) по току I_0 равны друг другу:

$$1 = (aUR_t) I_{01} e^{(aUR_t) I_0} = (aUR_t) I_{01}; \quad (2-59)$$

при этом показатель степени в правой части (2-58) равен единице и, значит, $I_0 = eI_{01}$. Подставляя это значение в [(2-45)], получаем напряжение теплового пробоя:

$$U_T = \frac{1}{aeR_t I_{01}} \approx \frac{3}{\varphi_s R_t I_{01}}. \quad (2-60)$$

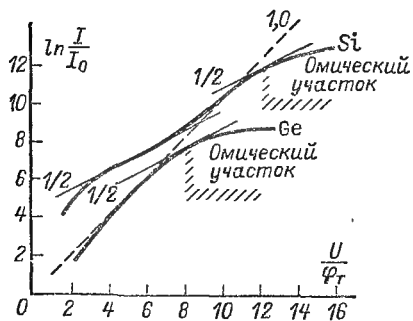
Например, если $\varphi_s = 0,7$ В; $R_t = 500^\circ \text{C/Вт}$ и $I_{01} = 20$ мкА, то $U_T \approx 400$ В. Значение U_T быстро уменьшается с ростом температуры окружающей среды, так как при этом сильно растет начальный ток I_{01} . У кремниевых приборов ток I_{01} настолько мал, что тепловой пробой практически исключается.

Заметим, что если обратный ток диода существенно возрос благодаря туннельному эффекту или ударной ионизации, то после этого может наступить тепловой пробой, так как напряжение U_T при большом токе I_{01} окажется меньше напряжения U_Z или U_M . Соответственно после вертикального участка, свойственного полемому и лавинному пробую, может иметь место участок с отрицательным сопротивлением.

2-8. ПРЯМАЯ ХАРАКТЕРИСТИКА РЕАЛЬНОГО ДИОДА

При напряжении $U > \varphi_T$ прямая ветвь характеристики согласно (2-33) должна быть экспоненциальной функцией, которая в полулогарифмическом масштабе представляется прямой линией с единичным наклоном (рис. 2-31, штриховая линия). Между тем, как видно из того же рисунка, реальные характеристики состоят из нескольких участков с разными наклонами, так что формула (2-33) представляет собой лишь некоторое приближение. Рассмотрим главные причины, по которым реальная характеристика отличается от идеализированной.

Рис. 2-31. Прямые характеристики реальных диодов — германиевого и кремниевых — в полулогарифмическом масштабе. Цифры характеризуют наклон кривых.



Ток рекомбинации. Из § 2-6 известно, что в равновесном состоянии токи термогенерации и рекомбинации в переходе вза-

имно компенсируются. При прямом смещении перехода крутизна потенциального барьера уменьшается и носители, не способные преодолеть барьер (см. рис. 2-6), проникают в переход гораздо

глубже. Соответственно увеличиваются их рекомбинации в переходе и ток рекомбинации I_R . Для оценки этого тока нужно найти скорость рекомбинации из общего выражения (1-59).

Обычно, как и при расчете тока I_G , считают уровни ловушек расположенными в середине запрещенной зоны, т. е. $n_t = p_t = n_i$.

Рассмотрим участок перехода, в котором концентрации электронов и дырок одинаковы (такой участок всегда имеется внутри перехода, см. рис. 2-9, а). Полагая $p = n$ в (2-18) и $U > 2\varphi_T$, получаем для этого участка:

$$p = n = n_i e^{U/2\varphi_T} \gg n_i = p_t = n_t;$$

$$pn = n_i^2 e^{U/\varphi_T} \gg n_i^2 = n_0 p_0.$$

Тогда скорость рекомбинации согласно (1-59б) имеет вид:

$$V \approx -\frac{n_i}{\tau_\infty} e^{U/2\varphi_T}.$$

Умножив модуль $|V|$ на заряд q и объем перехода Sl , можно получить приближенное значение тока рекомбинации. Более точное значение получится, если проинтегрировать величину V вдоль перехода, поскольку на разных его участках концентрации p и n различны и находятся в разных соотношениях с концентрациями n_t и p_t . Такой более строгий анализ [31] приводит к следующему выражению:

$$I_R = q(Sl) \frac{\varphi_T}{\Delta\varphi_0 - U} \frac{n_i}{\tau} e^{U/2\varphi_T}. \quad (2-61)$$

Как видим, ток I_R , подобно току I_G , пропорционален собственной концентрации n_i , а потому его значение и доля его в общем прямом токе диода существенно зависят от материала. Ток рекомбинации, так же как и ток термогенерации, играет главную роль в кремниевых диодах. В германиевых диодах его роль может стать заметной при пониженной температуре, когда тепловой ток I_0 сильно уменьшается.

Главная особенность выражения (2-61) состоит в том, что показатель степени экспоненты содержит множитель $1/2$, которого нет в формуле (2-33). Соответственно наклон такой характеристики в полулогарифмическом масштабе составляет $1/2$. Очевидно, что диффузионный ток (2-33) сильнее зависит от напряжения, чем ток рекомбинации (2-61). Поэтому даже тогда, когда ток I_R играет главную роль при малых напряжениях, с ростом напряжения он неизбежно уступает эту роль диффузионному току. В кремниевых диодах это имеет место при напряжении $U = 0,2 \div 0,3$ В, т. е. в пределах «пятки» (см. с. 119), что легко показать, приравняв значения токов, полученные по формулам (2-33) и (2-61).

Сопротивление базы. Будем считать, что в отсутствие инжекции или при малом ее уровне сопротивление базы определяется обычной формулой ¹:

$$r_6 = \rho_6 \frac{w}{S}, \quad (2-62)$$

где w — толщина базы; S — площадь поперечного сечения.

Например, при $\rho_6 = 5 \text{ Ом} \cdot \text{см}$; $w = 0,02 \text{ см}$; $S = 0,01 \text{ см}^2$ получим $r_6 = 10 \text{ Ом}$. При других параметрах базы сопротивление r_6 обычно лежит в пределах от 1—2 до 20—30 Ом. Падение напряжения на базе составляет:

$$U_6 = I r_6 = j \rho_6 w. \quad (2-63)$$

Это напряжение является той поправкой, которую, вообще говоря, следует ввести в формулы (2-33) и (2-37), чтобы учесть различие между падением напряжения на переходе U и приложенным напряжением $U_{\text{эб}}$:

$$I = I_0 \left(e^{\frac{U_{\text{эб}} - U_6}{\Phi_T}} - 1 \right); \quad (2-64a)$$

$$U_{\text{эб}} = \Phi_T \ln \left(\frac{I}{I_0} + 1 \right) + I r_6. \quad (2-64b)$$

В области малых токов поправка $I r_6$ мала и ею пренебрегают (см. п. 3, с. 114). Однако с увеличением тока напряжение U_6 растет линейно, а напряжение U — логарифмически, т. е. более слабо. Поэтому при достаточно большом токе всегда превалирует напряжение U_6 и экспоненциальная характеристика диода вырождается. Вырожденный участок, называемый *омическим*, в полулогарифмическом масштабе имеет меньший наклон (см. рис. 2-31). В линейном масштабе омический участок показан на рис. 2-32*.

Оценим значение тока, при котором наступает вырождение экспоненциальной характеристики. Для этого проще всего положить дифференциальное сопротивление r_d равным сопротивлению r_6 . Такое равенство соответствует одинаковым приращениям напряжений U и U_6 при изменении тока на ΔI . Используя (2-38), приходим к соотношению

$$I_v = \frac{\Phi_T}{r_6}, \quad (2-65)$$

¹ Формула (2-62) соответствует формуле (1-106), т. е. подразумевает однородность и одномерность базы (рис. 2-1, а). На практике оба эти условия, особенно второе, часто не выполняются; тогда вместо (2-62) нужно было бы использовать формулу (1-107), однако при этом количественные расчеты сильно осложнились бы. Заметим еще, что наличие демберовского поля [см. (1-120)] означает повышенное падение напряжения вблизи перехода; это равносильно некоторому увеличению сопротивления r_6 по сравнению со значением этого сопротивления, найденным по формуле (2-62).

* Ниже будет показано, что вырожденный участок характеристики отнюдь не линеен (из-за модуляции сопротивления базы). Поэтому термин «омический» является условным и отражает лишь тот факт, что при достаточно больших токах зависимость тока от напряжения перестает быть экспоненциальной и становится значительно более слабой.

которое является критерием вырождения. Например, при $r_6 = 10 \text{ Ом}$ получаем ток вырождения $I_B = 2,5 \text{ мА}$. При токе I_B падение напряжения в базе $U_6 = I_B r_6 = \varphi_T$.

Даже у мощных диодов, имеющих сопротивление базы около 1 Ом , ток вырождения не превышает $20\text{—}30 \text{ мА}$, т. е. соответствует начальному участку характеристики (рис. 2-32). Отсюда следует, что омический участок может составлять значительную, а иногда и основную рабочую часть характеристики.

Зависимость прямого напряжения от температуры. Рассмотрим сначала идеализированный диод, пользуясь формулой (2-37), в которой от температуры зависят параметры φ_T и I_0 . Зависимость $U(T)$ удобно характеризовать разностью $\Delta U = U(T) - U(T_0)$, которая оказывается пропорциональной температуре: $\Delta U = \varepsilon \Delta T$, где температурная чувствительность напряжения ε — почти постоянная величина. Пренебрегая единицей в формуле (2-37), подставляя I_0 из выражения (2-43а) и логарифмируя, получаем:

$$U = \varphi_T \ln \frac{I}{I_{00}} + \varphi_a.$$

Здесь $\varphi_T \sim T$, а ток I_{00} и потенциал φ_a слабо зависят от температуры. Следовательно, функция $U(T)$ действительно близка к линейной. Полагая $\varphi_a = \text{const}$; $\ln(I/I_{00}) = \text{const}$ и дифференцируя U по T , легко получить температурную чувствительность напряжения перехода в следующем виде:

$$\varepsilon_n = \frac{dU}{dT} = \frac{U - \varphi_a}{T} < 0. \quad (2-66)$$

Зависимость температурной чувствительности от температуры в формуле (2-66) является кажущейся¹, так как с изменением температуры соответственно изменяется значение U . Отрицательный знак температурной чувствительности обусловлен очевидным соотношением

$$U < \Delta\varphi_0 < \varphi_a.$$

Модуль температурной чувствительности, как следует из (2-66), растет с уменьшением прямого напряжения, а значит, и прямого тока. В пределе, при $U \rightarrow 0$, модуль ε_n имеет максимальное значение φ_a/T . Наоборот, с увеличением тока модуль температурной чувствительности уменьшается и в пределе, при $U \rightarrow \Delta\varphi_0$, имеет мини-

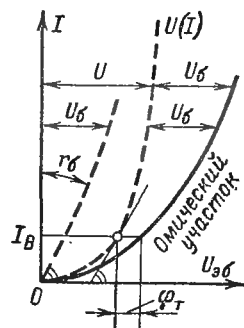


Рис. 2-32. Влияние объемного сопротивления базы на прямую характеристику реального диода.

¹ Аналогичный случай имел место при анализе температурной зависимости высоты равновесного потенциального барьера [см. формулу (2-5а) и замечание к ней]. Если учесть температурную зависимость коэффициента диффузии, входящего в (2-36в), то в числитель (2-66) добавляется член $-(c-1)\varphi_T$ [ср. с (2-56)]. Дальнейшее уточнение может быть связано с учетом зависимости $L(T)$ в (2-36в), однако обе поправки практически не очень существенны.

маленькое значение $(\varphi_s - \Delta\varphi_0)/T$. Для кремния и германия максимальные температурные чувствительности получаются разными (соответственно около 3 и около 2 мВ/°С), а минимальные температурные чувствительности — почти одинаковыми (около 1,2 мВ/°С). Последнее объясняется тем, что напряжения φ_s и $\Delta\varphi_0$ у германиевых и кремниевых диодов различаются примерно на одну и ту же величину 0,4 В. Можно считать типичным значением температурной чувствительности для обоих типов диодов $|\varepsilon_n| \approx 1,6$ мВ/°С.

Опыт показывает, что с ростом тока температурная чувствительность уменьшается сильнее, чем следует из (2-66). Более того, при некотором достаточно большом токе температурная чувствительность делается равной нулю и далее меняет знак, т. е. становится положительной [37]. Такие отклонения от формулы (2-66) объясняются наличием сопротивления базы, которое увеличивается с ростом температуры (см. рис. 1-22). Таким образом, температурная чувствительность, вообще говоря, определяется двумя приращениями (ΔU и ΔU_6), имеющими разные знаки. Роль приращения ΔU_6 становится определяющей на омическом участке характеристики, т. е. при достаточно больших токах, которые, однако, лежат в рабочем диапазоне диода.

Оценим температурную чувствительность, обусловленную базовым сопротивлением. Для этого продифференцируем функцию (1-366) по T , учитывая зависимость (1-32). В результате получим:

$$\frac{d\rho_6}{dT} = c \frac{\rho_6(T)}{T}.$$

Поскольку при заданном токе имеют место соотношения $dU_6 = I dr_6$ и $dr_6 \sim d\rho_6$, легко прийти к следующему выражению для температурной чувствительности напряжения базы:

$$\varepsilon_6 = \frac{dU_6}{dT} = c \frac{I r_6(T)}{T}. \quad (2-67)$$

Например, если $c = 3/2$, $I = 100$ мА, $r_6 = 10$ Ом и $T = 300$ К, то $\varepsilon_6 = +5$ мВ/°С. С изменением температуры ε_6 несколько меняется. Приравнивая значения $|\varepsilon_n|$ и ε_6 , можно найти ток I_e , при котором суммарная температурная чувствительность меняет знак:

$$I_e = \frac{1}{c} |\varepsilon_n| T / r_6.$$

Если $c = 3/2$, $|\varepsilon_n| = 1,6$ мВ/°С, $r_6 = 10$ Ом и $T = 300$ К, то $I_e \approx 35$ мА.

На рис. 2-33 показаны прямые характеристики кремниевого диода при трех температурах. Как видим, эти характеристики пересекаются, что происходит вследствие изменения значения и знака температурной чувствительности по мере возрастания прямого тока.

Работа диода при высоком уровне инжекции. Формула (2-33) была выведена из условия низкого уровня инжекции, при котором можно пренебречь дрейфовой составляющей тока инжектированных

носителей (см. п. 2, с. 113). Кроме того, низкий уровень инжекции означает, что удельное сопротивление базы вблизи перехода сохраняет исходное значение, поскольку концентрации избыточных носителей относительно малы,

Количественной характеристикой уровня инжекции (см. с. 81) является отношение

$$\delta = \frac{\Delta p(0)}{n_0},$$

где $\Delta p(0)$ — граничная концентрация избыточных дырок; n_0 — равновесная концентрация электронов в базе.

В условиях квазинейтральности, когда $\Delta n \approx \Delta p$, низкий уровень инжекции ($\delta \ll 1$) означает, что $\Delta n \ll n_0$, т. е. распределение электронов в базе почти не нарушено.

С ростом прямого тока и напряжения уровень инжекции повышается, растут избыточные концентрации дырок и компенсирующих электронов и процессы в базе существенно изменяются.

Оценим значение тока, до которого действительны диффузионное приближение и, следовательно, функция (2-33). Сравнивая (2-28а) и (2-33), получаем:

$$\Delta p(0) = p_0 \frac{I}{I_0}. \quad (2-68)$$

Деля обе части (2-68) на n_0 , можно представить уровень инжекции в следующей форме:

$$\delta = \frac{\Delta p(0)}{n_0} = \frac{I}{I_\delta}, \quad (2-69)$$

где $I_\delta = I_0 (n_0/p_0)$ — тот «масштабный» ток, относительно которого оценивается уровень инжекции¹. Назовем I_δ *граничным током*.

Используя формулы (2-36а), (1-74) и (1-36б), получаем:

$$I_\delta = \sigma_6 \frac{\Phi_T}{bL \operatorname{th}(w/L)} S. \quad (2-70а)$$

В случае толстой базы, когда $\operatorname{th}(w/L) \approx 1$,

$$I_\delta = \sigma_6 \frac{\Phi_T}{bL} S. \quad (2-70б)$$

В случае тонкой базы [когда $\operatorname{th}(w/L) \approx w/L$]

$$I_\delta = \sigma_6 \frac{\Phi_T}{bw} S. \quad (2-70в)$$

¹ Учитывая соотношения (1-16) и (1-36), можно установить следующую связь между токами I_δ и I_0 :

$$I_\delta = \left(1 + \frac{1}{b}\right)^2 \left(\frac{\rho_i}{\rho_6}\right)^2 I_0 \approx 2 \left(\frac{\rho_i}{\rho_6}\right)^2 I_0,$$

где коэффициент b (1-31) для германия и кремния близок к 2,5.

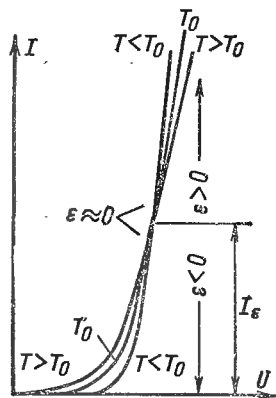


Рис. 2-33. Температурная зависимость прямой характеристики реального диода.

Для типичных значений $\rho_0 = (2 \div 10) \text{ Ом} \cdot \text{см}$ и $\omega \approx L = 100 \text{ мкм}$ ток I_0 (при площади $S = 0,01 \text{ см}^2$) составляет 1—6 мА. В этих пределах (при данной площади) действительно выражение (2-33).

Сравнивая (2-65) и (2-70в), нетрудно заметить, что токи I_v и I_0 близки друг к другу, т. е. вырождение экспоненциальной характеристики в омическую совпадает с переходом от низкого к высокому уровню инжекции.

Особенности, связанные с высоким уровнем инжекции (когда $\delta \gtrsim 1$), следующие.

1. Концентрация дырок вблизи перехода становится сравнимой с концентрацией электронов. В связи с этим на участке, протяженность которого примерно равна диффузионной длине, заметную роль начинает играть дрейфовая составляющая дырочного тока.

2. Увеличение избыточных концентраций Δp и Δn вблизи перехода приводит к существенному увеличению удельной проводимости базы в этой области (назовем этот участок *модулированным*). Модуляция сопротивления базы сказывается на коэффициенте инжекции (2-35) и на омическом падении напряжения U_0 (2-63).

Дрейфовая составляющая тока инжектированных носителей. Полагая коэффициент инжекции близким к единице, будем по-прежнему считать, что полный электронный ток на эмиттерной границе базы значительно меньше дырочного; в первом приближении можно положить $j_n = 0$. Тогда напряженность поля в базе будет выражаться формулой (1-91). Подставляя это значение E в (1-72а) и учитывая соотношение (1-74), получаем дрейфовую составляющую дырочного тока:

$$(j_p)_{др} = -qD_p \frac{p}{n} \frac{dn}{dx}.$$

Если использовать условия квазинейтральности (1-50), то производную dn/dx можно заменить на dp/dx , а концентрацию n записать как $n_0 + \Delta p$; кроме того, положим $p = \Delta p$, поскольку $\Delta p \gg p_0$. Тогда

$$(j_p)_{др} = -q \left(D_p \frac{\Delta p}{\Delta p + n_0} \right) \frac{dp}{dx}. \quad (2-71)$$

Как видим, дрейфовую составляющую удалось привести к такой же форме, как и диффузионную (1-73а). Складывая обе составляющие, получаем полный дырочный ток вблизи перехода:

$$j_p = -q \left[D_p \left(1 + \frac{\Delta p}{\Delta p + n_0} \right) \right] \frac{dp}{dx}. \quad (2-72)$$

Здесь произведение в квадратных скобках можно считать эквивалентным коэффициентом диффузии, значение которого зависит от уровня инжекции. При высоком уровне инжекции ($\Delta p(0) \gg n_0$) эквивалентный коэффициент диффузии близок к $2D_p$. Следовательно, при высоком уровне инжекции дрейфовая составляющая дырочного тока почти равна диффузионной составляющей, а значит, полный дырочный ток можно рассчиты-

вать по «диффузионной» формуле (2-33), удваивая получаемые значения:

$$I \approx 2I_0 e^{\frac{U}{\varphi_T}}. \quad (2-73)$$

Как видим, учет дрейфовой составляющей не приводит к качественным изменениям вольт-амперной характеристики диода.

Коэффициент инжекции. Как уже отмечалось, высокий уровень инжекции приводит к модуляции (уменьшению) удельного сопротивления базы.

Важным следствием модуляции базового сопротивления является уменьшение коэффициента инжекции, характеризующего долю дырочного тока на базовой границе перехода. Запишем граничную проводимость базы (при $x = 0$) в виде

$$\sigma_6(0) = 1/\rho_6 = \sigma_{60} + \Delta\sigma_6(0) = \sigma_{60} \left(1 + \frac{\Delta\sigma_6(0)}{\sigma_{60}} \right),$$

где исходная проводимость базового слоя

$$\sigma_{60} = 1/\rho_{60} = q\mu_n n_0,$$

а приращение проводимости (с учетом квазинейтральности)

$$\Delta\sigma_6(0) = q\mu_n \Delta n(0) + q\mu_p \Delta p(0) \approx q\mu_n \left(1 + \frac{1}{b} \right) \Delta p(0).$$

Для простоты положим $1/b \ll 1$; тогда $\Delta\sigma_6(0)/\sigma_{60} \approx \delta$ и, следовательно, согласно (2-35)

$$\gamma \approx 1 - \frac{\rho_3}{\rho_{60}} (1 + \delta). \quad (2-74)$$

Такое приближение, столь же грубое, как и (2-35), действительно для значений $\gamma \geq 0,95 \div 0,9$. Например, если $\rho_3/\rho_{60} = 0,001$, то формула (2-74) действительна для $\delta \leq 50 \div 100$, т. е. для весьма высоких уровней инжекции.

Модуляция сопротивления базы. Повышенная концентрация носителей, свойственная высокому уровню инжекции, влияет не только на коэффициент инжекции (см. предыдущий раздел), но и на полное сопротивление базы r_6 . Тем самым согласно (2-64) форма вольт-амперной характеристики зависит от уровня инжекции более сложно и более существенно, чем если бы эта зависимость определялась только увеличением эквивалентного коэффициента диффузии [см. (2-72)].

Для того чтобы найти функцию $r_6(\delta)$, запишем сначала сопротивление элементарного слоя базы толщиной dx , расположенного вблизи перехода:

$$dr_6 = \frac{dx}{(\sigma_{60} + \Delta\sigma_6)} S,$$

где (с учетом квазинейтральности, т. е. условия $\Delta n \approx \Delta p$)

$$\Delta\sigma_6 = q\mu_p \Delta p(x) + q\mu_n \Delta n(x) \approx q\mu_n \left(1 + \frac{1}{b} \right) \Delta p(x).$$

Подставляя в выражение для $\Delta\sigma_6$ распределение (2-30) и интегрируя dr_6 в пределах от $x = 0$ до $x = w$, можно привести сопротивление базы к следующему виду¹:

$$r_6(\delta) = \frac{r_{60}}{\frac{w}{L} \sqrt{1+k^2}} \ln \frac{e^{\frac{w}{L}} (1+k+\sqrt{1+k^2}) + (1-k-\sqrt{1+k^2})}{e^{\frac{w}{L}} (1+k-\sqrt{1+k^2}) + (1-k+\sqrt{1+k^2})}, \quad (2-75)$$

где

$$k = \frac{(b+1)/b}{\operatorname{sh}\left(\frac{w}{L}\right)} \delta \approx \frac{\delta}{\operatorname{sh}\left(\frac{w}{L}\right)}, \quad (2-76)$$

а r_{60} — сопротивление базы при низком уровне инжекции (2-62).

Графики функции $r_6(\delta)$ приведены на рис. 2-34 в относительном масштабе. Конечно, в общем виде зависимость (2-75) слишком сложна для расчетов. Поэтому рассмотрим два частных случая, в которых эта зависимость существенно упрощается.

Пусть $k \ll 1$ (практически достаточно, если $k < 0,5$). Этот случай соответствует не очень высоким уровням инжекции или достаточно толстой базе. Тогда в формуле (2-75) можно положить $k^2 = 0$, а также пренебречь величиной k везде, кроме первого члена знаменателя. Этот первый член ($ke^{w/L}$) с учетом (2-76) будет примерно равен 2δ (по-

скольку $\operatorname{sh}(w/L) \approx 0,5e^{w/L}$, если $w/L \gg 1$). В результате после логарифмирования дроби получаем²:

$$r_6(\delta) \approx r_{60} \left[1 - \frac{L}{w} \ln(1+\delta) \right]. \quad (2-77a)$$

На рис. 2-34 зависимость (2-77a) действительна для начальных участков кривых вплоть до точек, отмеченных знаком $\times$.

Пусть теперь $k \gg 1$ (практически достаточно, если $k > 2$). Этот случай соответствует высоким уровням инжекции или достаточно тонкой базе. Тогда в формуле (2-75) можно положить

¹ При интегрировании используется подстановка $y = e^{\frac{w-x}{L}}$ с соответствующей заменой пределов интегрирования.

² Это выражение в несколько иной форме, а главное, без указания границ применимости приведено в [38].

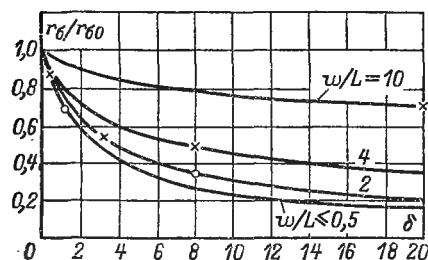


Рис. 2-34. Зависимость базового сопротивления от уровня инжекции и толщины базы.

$\sqrt{1+k^2}=k$ и привести ее к следующему виду:

$$r_6(\delta) \approx r_{60} \frac{\operatorname{sh} \frac{w}{L}}{\frac{w}{L} \delta} \ln \left(1 + \frac{\delta}{\operatorname{ch}^2(w/2L)} \right). \quad (2-776)$$

На рис. 2-34 зависимость (2-776) действительна для конечных участков кривых, начиная от точек, отмеченных знаком $\circ$. Случай $k \gg 1$ характерен для реальных диодов. Поэтому упростим выражение (2-776) еще более, полагая $\operatorname{sh}(w/L) \approx w/L$, $\operatorname{ch}(w/2L) \approx 1$ и $\delta \gg 1$:

$$r_6(\delta) \approx \frac{r_{60}}{\delta} \ln \delta. \quad (2-77в)$$

Теперь посмотрим, как отражается модуляция базового сопротивления на форме вольт-амперной характеристики диода¹. Для этого запишем сначала падение напряжения на базе. Используя выражение (2-776) и подставляя в него (2-62), (2-69) и (2-70а), получаем:

$$U_6 = I r_6 = m' \varphi_T \ln [1 + \delta / \operatorname{ch}^2(w/2L)], \quad (2-78а)$$

где

$$m' = \frac{1}{b} \operatorname{ch}(w/L).$$

Если пренебречь единицей под логарифмом и учесть (2-69), нетрудно записать вольт-амперную характеристику модулированного базового слоя в виде

$$I = I_0 \operatorname{ch}^2 \left(\frac{w}{2L} \right) e^{\frac{U_6}{m' \varphi_T}}. \quad (2-78б)$$

Выражения (2-78) подтверждают, что вырожденный («омический») участок характеристики нелинеен (см. сноску на с. 136 и рис. 2-32).

Чтобы получить вольт-амперную характеристику диода в целом, нужно подставить (2-78а) в выражение (2-64б) и раскрыть значение δ согласно (2-69). Тогда после небольших преобразований получаем:

$$U_{96} = \varphi_T \ln \left[\left(\frac{I}{2I_0} + 1 \right) \left(\frac{I}{I_0 \operatorname{ch}^2(w/2L)} + 1 \right)^{m'} \right], \quad (2-79а)$$

где коэффициент 2 при токе I_0 отражает роль дрейфовой составляющей [см. (2-73)].

Пренебрегая единицами в круглых скобках (2-79а) и выражая ток как функцию напряжения U_{96} , получаем вольт-амперную ха-

¹ Более подробно этот вопрос рассмотрен в работах [39, 40].

рактеристику диода к такой же форме, как (2-73):

$$I = 2hI_0 e^{\frac{U_{\text{эб}}}{m\varphi_T}}. \quad (2-796)$$

Здесь приняты следующие обозначения:

$$h = \left[\frac{I_8}{2I_0} \operatorname{ch}^2(\omega/2L) \right]^{\frac{m-1}{m}};$$

$$m = 1 + m' = 1 + \frac{\operatorname{ch}(\omega/L)}{b}.$$

При $b = 2$ типичны значения $m = 1,5 \div 2$; они соответствуют значениям $0 < \omega/L \leq 1,5^*$. При этом наклон логарифмической кривой (см. рис. 2-31) составляет $2/3 - 1/2$. У кремниевых диодов, у которых начальный участок характеристики описывается формулой (2-61), наклон, близкий к $1/2$, обычно получается на всем протяжении характеристики.

До сих пор специфика высокого уровня инжекции рассматривалась с позиций уточнения формулы (2-33). Поэтому результаты уточнения в виде формулы (2-796) не изменили экспоненциальный характер зависимости тока от напряжения. Напомним, что такая зависимость является следствием распределения Больцмана, которое сохраняется до тех пор, пока область p - n перехода находится в квазиравновесном состоянии (см. с. 100). Естественно, что с ростом тока условие (2-17) рано или поздно нарушается и тогда соотношения (2-13), а значит, и граничное условие (2-28а) теряют силу. Соответственно при достаточно большом токе будут принципиально недействительны не только простейшие формулы (2-33) и (2-73), но и уточненная формула (2-796).

Такие качественные изменения наступают тогда, когда высота потенциального барьера уменьшается примерно до φ_T , а граничная концентрация дырок в базе $p_6(0)$ делается сравнимой с равновесной концентрацией дырок в эмиттере p_{30} . При этом согласно (2-74) коэффициент инжекции становится заметно меньше единицы ($\gamma < 0,9$). Анализ показывает [41], что в области таких сверхвысоких уровней инжекции (когда $\delta > 0,1 p_{30}/n_{60}$) вольт-амперная характеристика перестает быть экспоненциальной, постепенно приближаясь к степенной. С некоторыми допущениями можно получить для этого режима квадратичную зависимость:

$$I = aU^2, \quad (2-80)$$

где коэффициент a зависит от параметров материала базы и обратно пропорционален ее толщине в степени $2/3$. Функция (2-80) несравненно более слабая, чем (2-796); в полулогарифмическом масштабе (см. рис. 2-31) она изображается почти горизонтальной линией.

Распределение токов в базе. Из рис. 2-35 и 2-36 видно, что соотношение между отдельными компонентами тока во многом зависит от уровня инжекции и относительной толщины базы ω/L . Кри-

* Точное значение $m = 2$ получено в обзоре [26] для диода типа p - i - n (с промежуточным слоем собственного полупроводника). В этом случае происхождение коэффициента m иное: оно связано с распределением приложенного напряжения между двумя переходами, так что на каждом из них действует половина напряжения. Другие случаи, когда $m > 1$, рассмотрены в том же обзоре.

вые на этих рисунках подтверждают предпосылки, которые делались при анализе диода, а также позволяют сделать некоторые общие выводы.

1. Ток диода, неизменный вдоль оси x , складывается из полных токов основных и неосновных носителей, из которых первый возрастает, а второй убывает по мере удаления от перехода.

2. Если коэффициент инжекции близок к единице, то полный ток диода независимо от уровня инжекции определяется полным током неосновных носителей на границе перехода (в данном случае — граничным током дырок).

3. В соответствии с почти одинаковым распределением избыточных носителей диффузионные составляющие электронного и дырочного токов находятся в соотношении $D_n : D_p$ и противоположны по направлению.

4. Распределение дрейфовых составляющих токов сложнее, чем диффузионных: дрейфовые составляющие находятся в соотношении $n\mu_n / p\mu_p$, которое зависит и от уровня инжекции, и от координаты. Кроме того, распределение дрейфовых составляющих зависит от распределения поля: при низких уровнях инжекции поле максимально вблизи перехода [благодаря эффекту Дембера, см. (1-118)], а при высоких уровнях инжекции — вдали от перехода, поскольку начальный участок базы модулирован и имеет пониженное удельное сопротивление (см. предыдущий раздел).

5. Для толстой базы (рис. 2-35, а и 2-36, а) ток основных носителей (в данном случае электронов) мал только вблизи перехода; на расстоянии $(2 \div 3) L$ от перехода, где слой базы находится почти

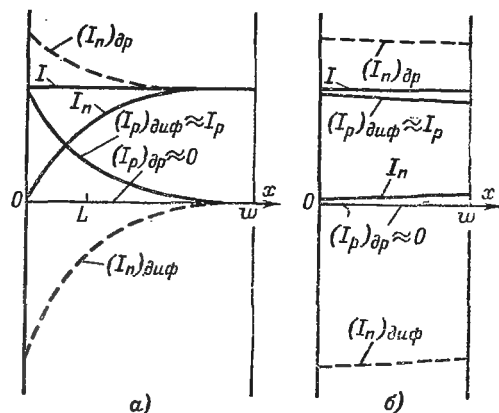


Рис. 2-35. Распределение токов в базе диода при низком уровне инжекции.

а — толстая база; б — тонкая база.

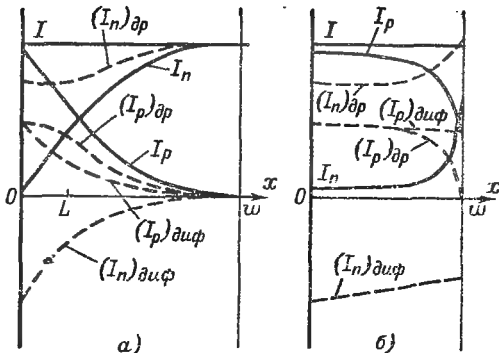


Рис. 2-36. Распределение токов в базе диода при высоком уровне инжекции.

а — толстая база; б — тонкая база.

в равновесном состоянии и роль неосновных носителей ничтожна, ток основных носителей является главным компонентом и имеет практически чисто дрейфовый характер, как в обычном однородном полупроводнике.

6. Для толстой базы электронная и дырочная составляющие полного тока становятся одинаковыми на расстоянии, примерно равном диффузионной длине от перехода. Для тонкой базы такое равенство либо имеет место только вблизи омического контакта

(при высоком уровне инжекции, см. рис. 2-36, б), либо вообще не имеет места (при низком уровне инжекции, когда полный ток диода определяется током неосновных носителей, см. рис. 2-35, б).

В заключение следует отметить, что ни в предшествующем анализе, ни на рис. 2-35 и 2-36 не учитывались зависимости параметров μ , D и τ от концентраций подвижных носителей, что, конечно, не может не отразиться на точности анализа.

Эквивалентная схема диода при прямом смещении. Несмотря на сложную структуру прямой характеристики, для практических расчетов ее можно аппроксимировать ломаной линией, как показано на рис. 2-37, а. Погрешность такой аппроксимации, как видно из рисунка, существенна лишь на начальном участке, при малых токах. Формула для идеализированной характеристики очень проста:

$$I = \frac{U - E_{np}}{r_{np}}. \quad (2-81)$$

Ей отвечает простая эквивалентная схема, показанная на рис. 2-37, б. Параметры r_{np} и E_{np} можно определить либо по справочным (типовым), либо по специально снятым (индивидуальным) кривым. Напомним, что у кремниевых диодов (см. рис. 2-21, б) значение E_{np} выражено значительно ярче, чем у германиевых. Соответственно точнее оказывается аппроксимация (2-81). В диапазоне 0,5—5 мА для кремниевых диодов принимают $E_{np} \approx 0,7$ В.

2-9. ПЕРЕХОДНЫЕ ХАРАКТЕРИСТИКИ ДИОДА

Полупроводниковый диод является инерционным элементом по отношению к быстрым изменениям тока или напряжения, поскольку новое распределение носителей устанавливается не сразу. Как известно, внешнее напряжение меняет ширину перехода (см. § 2-2 и 2-3), а значит, и величину пространственных зарядов в переходе. Кроме того, при инжекции (или экстракции)

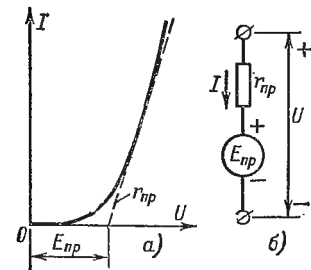


Рис. 2-37. Прямая характеристика реального диода, ее идеализация (а) и эквивалентная схема диода при прямом включении (б).

меняются заряды в квазиинейной области базы. Следовательно, наряду с проводимостью, которая в первом приближении характеризуется выражением (2-33), диод обладает емкостью, которую можно считать подключенной параллельно р-п переходу. Эту емкость принято разделять на две составляющие: барьерную емкость, отражающую перераспределение зарядов в переходе, и диффузионную емкость, отражающую перераспределение зарядов в базе. Такое разделение в общем весьма условно, но удобно на практике, тем более что соотношение обеих емкостей различно при разных полярностях смещения. При прямом смещении главную роль играют заряды в базе и соответственно диффузионная емкость. При обратном смещении (режим экстракции) заряды в базе меняются мало и главную роль играет барьерная емкость.

Характеристика инерционных свойств диода с помощью емкостей особенно удобна в случае малых переменных сигналов, действующих на фоне больших постоянных смещений. При этом емкости оказываются почти линейными и весьма наглядно дополняют эквивалентную схему диода. В случае больших сигналов использование емкостей, особенно диффузионной, становится нецелесообразным, так как их нелинейность проявляется в такой мере, что эквивалентная схема теряет свою наглядность, а анализ делается отнюдь не более простым, чем при использовании уравнений непрерывности.

После этих предварительных замечаний мы рассмотрим сначала свойства барьерной и диффузионной емкостей, а затем исследуем переходный процесс при большом сигнале.

Барьерная емкость (емкость перехода). Определим величину барьерной емкости, считая, что переход несимметричен и сосредоточен в базе типа n (т. е. $l \approx l_n$).

Заряды на пластинах воображаемого конденсатора — перехода — одинаковы ($-Q_p = Q_n$). Найдем один из них, например Q_n , для случая ступенчатого перехода (см. рис. 2-3, в):

$$Q_n = qN_d(Sl),$$

где S — площадь перехода; l — ширина перехода.

Подставляя сюда выражение (2-12) и дифференцируя Q_n по $|U|$, получаем дифференциальную барьерную емкость ступенчатого перехода при обратном смещении $|U| \gg \Delta\phi_0$:

$$C_n = \frac{S \sqrt{0,5 \epsilon_0 e q N_d}}{\sqrt{|U|}} = \frac{\epsilon_0 e S}{l_0} \sqrt{\frac{\Delta\phi_0}{|U|}}. \quad (2-82a)$$

Вторая форма выражения удобна тем, что ее первый множитель является емкостью обычного плоского конденсатора с расстоянием между обкладками, равным l_0 — равновесной ширине перехода, определяемой по формуле (2-96). Как видим, емкость перехода представляется несколько «необычной», потому что ее величина зависит от приложенного напряжения. Оценим значение барьерной емкости

для кремниевого диода, полагаая $l_0 = 0,5$ мкм; $\Delta\varphi_0 = 0,75$ В; $S = 0,01$ см² и $|U| = 20$ В; тогда $C_{\pi} \approx 50$ пФ.

В общем случае, используя (2-11), для барьерной емкости будем иметь:

$$C_{\pi} = \frac{\epsilon_0 \epsilon S}{l_0} \sqrt{\frac{\Delta\varphi_0}{\Delta\varphi_0 - U}}. \quad (2-826)$$

Эта формула согласно [42] может давать большую погрешность при прямых напряжениях более 0,1—0,2 В, так как в исходных выражениях (2-9) не учтен заряд подвижных носителей в переходе, существенный при прямых смещениях. Необходимо заметить, что изменение ширины перехода (т. е. перезаряд емкости) требует притока или отвода основных (для каждого слоя) носителей. Отсюда следует, что на высоких частотах уменьшается коэффициент инжекции неосновных носителей (с. 118). Точнее говоря, коэффициент инжекции в области высоких частот становится комплексной величиной [см. (4-30)].

В случае узкого плавного перехода, смещенного в обратном направлении, барьерная емкость получается тем же методом, но с использованием выражения (2-236):

$$C_{\pi} = \frac{S \sqrt[3]{0,11 (\epsilon_0 \epsilon)^2 q N'}}{\sqrt[3]{|U|}} = \frac{\epsilon_0 \epsilon S}{l_0} \sqrt[3]{\frac{\Delta\varphi_0}{|U|}}, \quad (2-83)$$

где N' — градиент концентрации в области перехода.

Зависимость барьерных емкостей от обратного напряжения показана на рис. 2-38. Заметим, что емкость перехода при прочих равных условиях зависит от концентрации примесей, т. е. согласно (1-36) от удельного сопротивления материала. Чем больше удельное сопротивление слоев, тем меньше емкость.

Иногда барьерную емкость определяют как интегральную величину $Q/|U|$. Легко убедиться, что выражения (2-82) и (2-83) при этом умножаются соответственно на 2 и на $3/2$.

Диффузионная емкость. При прямом смещении диода переход, как известно, сужается, и соответственно растет барьерная емкость [см. (2-826)]. Однако она оказывается менее существенной, чем емкость, обусловленная возрастающим зарядом носителей в базовом слое, которую называют диффузионной, так как этот же самый заряд лежит в основе диффузии носителей в базе.

Диффузионная емкость C_d «заряжается» как инжесктированными дырками, так и электронами, компенсирующими заряд инжек-

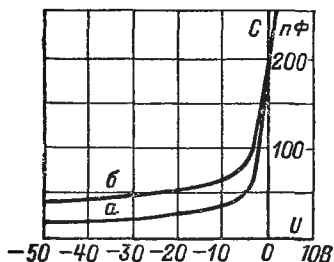


Рис. 2-38. Зависимость емкости ступенчатого (а) и плавного (б) переходов от обратного напряжения.

тированных дырок. Поскольку из бытовых заряды электронов и дырок одинаковы, найдем один из них, а именно заряд дырок, исходя из распределения (2-30):

$$\Delta Q = qS \int_0^w \Delta p(x) dx = \frac{qSL\Delta p(0)}{\text{th}(w/L)} \left(1 - \text{sech} \frac{w}{L}\right).$$

Подставляя сюда $\Delta p(0)$ из (2-68), ток I_0 из (2-36а) и учитывая соотношение (1-111), получаем:

$$\Delta Q = I\tau \left(1 - \text{sech} \frac{w}{L}\right). \quad (2-84)$$

Если поделить этот заряд на напряжение U , то интегральная диффузионная емкость запишется в следующем виде:

$$\bar{C}_d = \frac{\Delta Q}{U} = \frac{\tau}{R_d} \left(1 - \text{sech} \frac{w}{L}\right),$$

где $R_d = U/I$ — сопротивление диода постоянному току (2-39). Дифференциальная диффузионная емкость будет иметь аналогичную форму:

$$C_d = \frac{d(\Delta Q)}{dU} = \frac{\tau}{r_d} \left(1 - \text{sech} \frac{w}{L}\right) = \frac{I\tau}{\Phi_T} \left(1 - \text{sech} \frac{w}{L}\right), \quad (2-85)$$

где $r_d = dU/dI$ — сопротивление диода переменному току (2-38). Как видим, диффузионная емкость (2-85) является функцией прямого тока, подобно тому как барьерная емкость (2-82а) является функцией обратного напряжения. Кроме того, диффузионная емкость находится в прямой зависимости от толщины базы, уменьшаясь с уменьшением отношения w/L .

Для толстой базы, когда $w \gg L$ и $\text{sech}(w/L) \approx 0$, получаем:

$$\Delta Q \approx I\tau; \quad (2-86а)$$

$$C_d \approx \frac{\tau}{r_d} \approx \frac{I\tau}{\Phi_T}. \quad (2-86б)$$

Например, если $\tau = 5$ мкс; $I = 10$ мА, то $C_d = 2$ мкФ. Такие значения на несколько порядков превосходят величину барьерной емкости.

Для тонкой базы, когда $w < L$ и $\text{sech}(w/L) \approx 1 - 0,5(w/L)^2$, выражения для ΔQ и C_d приводятся к виду

$$\Delta Q \approx It_D; \quad (2-87а)$$

$$C_d \approx \frac{t_D}{r_d} \approx \frac{It_D}{\Phi_T}, \quad (2-87б)$$

где

$$t_D = \frac{w^2}{2D} \quad (2-88)$$

есть среднее время диффузии, т. е. среднее время пролета носителей через тонкую базу при диффузионном механизме движения¹.

Время диффузии является столь же фундаментальным параметром полупроводниковых приборов в случае тонкой базы, как время жизни для толстой базы. Например, сравнивая формулы (2-86) и (2-87), видим, что они имеют одинаковую структуру и различаются только тем, что место параметра τ в первых занимает параметр t_D во вторых. Поскольку случай тонкой базы характерен для транзисторов и многих других приборов, их динамические параметры в решающей степени определяются именно временем диффузии, даже если наряду с диффузией имеет место дрейф.

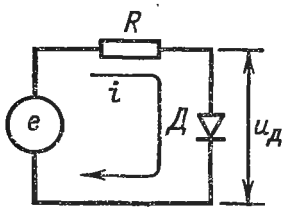


Рис. 2-39. Схема переключения диода.

Режим переключения. При включении и выключении диода переходные процессы проявляются в наиболее полном виде. Определенную роль в этих процессах играет барьерная емкость, однако в дальнейшем ее влияние для простоты не учитывается.

Анализ переходных процессов проводится обычно для ступенчатого сигнала (рис. 2-39, 2-40), когда диод попеременно работает в прямом и обратном направлениях (в частном случае обратное переключение может отсутствовать, т. е. диод просто отключается: $e_2 = 0$). Имеющиеся решения этой задачи (см. [43, 44] и библиографию в них) с математической точки зрения достаточно строги, но слишком сложны для инженерной практики. Кроме того, исходные предпосылки в любом случае являются лишь некоторым приближением по отношению к процессам в реальных диодах. Поэтому полезно провести упрощенный анализ [45] и получить сравнительно простые выражения для длительностей отдельных этапов переходного процесса.

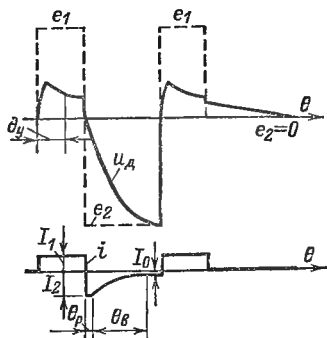


Рис. 2-40. Характер переходных процессов при отпирании, переключении и запираии диода.

При попеременном переключении диода с прямого направления на обратное различают следующие «искаженные» (динамические) участки переходной характеристики (рис. 2-40):

¹ В самом деле, поскольку в тонкой базе влияние рекомбинации слабо и распределение дырок почти линейно (см. рис. 2-19), диффузионный дырочный ток оказывается практически постоянным на всем протяжении базы (рис. 2-35, б). Значит, отношение избыточного заряда к этому току есть время перемещения заряда в базе, что соответствует выражению (2-87а).

1) установление прямого напряжения при заданном прямом токе (интервал θ_y);

2) рассасывание избыточных носителей в базе при заданном обратном токе (интервал θ_p);

3) восстановление обратного тока (сопротивления) при заданном обратном напряжении (интервал θ_n).

Рассмотрим последовательно эти участки.

Установление прямого напряжения. В подавляющем большинстве случаев на этом участке можно считать заданным прямой ток I_1 . Действительно, приложенное напряжение e_1 обычно много больше прямого падения напряжения на диоде U_d ; поэтому

$$I_1 = \frac{e_1 - U_d}{R} \approx \frac{e_1}{R} = \text{const.}$$

Прямое напряжение на диоде U_d состоит из двух компонентов: напряжения на p - n переходе U и напряжения на базовом слое U_b .

Определим первый компонент $u(t)$, воспользовавшись операторным диффузионным уравнением (1-121):

$$\frac{d^2(\Delta p)}{dx^2} - \frac{1 + s\tau}{L^2} \Delta p = 0. \quad (2-89)$$

Общим решением этого уравнения является функция

$$\Delta p(x) = A_1 e^{\frac{x}{L} \sqrt{1+s\tau}} + A_2 e^{-\frac{x}{L} \sqrt{1+s\tau}}. \quad (2-90)$$

Подставим сюда граничные условия для толстой базы:

$$\begin{aligned} -SqD_p \left| \frac{\partial(\Delta p)}{\partial x} \right|_{x=0} &= I_1; \\ \Delta p(\infty) &= 0, \end{aligned}$$

первое из которых в отличие от (2-28а) соответствует заданному току, а не напряжению. Тогда нетрудно получить изображение избыточной концентрации в следующем виде:

$$\Delta p(x; s) = p_0 \frac{I_1}{I_0 \sqrt{1+s\tau}} e^{-\frac{x}{L} \sqrt{1+s\tau}}.$$

Полагая $x = 0$, получаем изображение граничной концентрации:

$$\Delta p(0; s) = p_0 \frac{I_1}{I_0 \sqrt{1+s\tau}}, \quad (2-91)$$

затем по таблицам [22] находим оригинал:

$$\Delta p(0; \theta) = p_0 \frac{I_1}{I_0} \text{erf} \sqrt{\theta}, \quad (2-92)$$

где $\theta = t/\tau$ — относительное время, которым будем пользоваться в дальнейшем. Функция ошибок $\text{erf}(z)$ охарактеризована в сноске на с. 84.

Вместо функции ошибок $\operatorname{erf} \sqrt{\theta}$ можно, как правило, использовать простую аппроксимацию

$$\operatorname{erf} \sqrt{\theta} \approx \sqrt{1 - e^{-\theta}}, \quad (2-93a)$$

которая при малых и больших θ соответственно упрощается:

$$\operatorname{erf} \sqrt{\theta} \approx \sqrt{\theta} \quad \text{при } \theta < 0,5; \quad (2-93б)$$

$$\operatorname{erf} \sqrt{\theta} \approx 1 - \frac{1}{2} e^{-\theta} \quad \text{при } \theta > 0,5. \quad (2-93в)$$

Приравняв (2-92) выражению для неравновесной граничной концентрации (2-28a), легко найти искомую функцию $u(\theta)$:

$$u(\theta) = \varphi_T \ln \left(1 + \frac{I_1}{I_0} \operatorname{erf} \sqrt{\theta} \right). \quad (2-94)$$

Начальным значением этой функции является $U(0) = 0$, а конечным значением $U(\infty)$ — функция (2-37) при $I = I_1$.

Функция (2-94) имеет почти скачкообразный характер (рис. 2-41, а). Например, если $I_1/I_0 = 10^5$, то при $\theta = 0,01$ имеем уже $U = 0,8U(\infty)$. Время установления на уровне 0,9, которое обозначим θ'_y , нетрудно получить, полагая $U(\theta'_y) = 0,9U(\infty)$, пренебрегая единицей в выражении (2-94) и используя аппроксимацию (2-93a). После преобразования получим:

$$\theta'_y = \ln \frac{1}{1 - \left(\frac{I_1}{I_0} \right)^{-0,2}}. \quad (2-95)$$

Например, если $I_1/I_0 = 10^5$, то $\theta'_y = 0,1$.

Второй компонент прямого напряжения $u_6(t)$ можно найти,

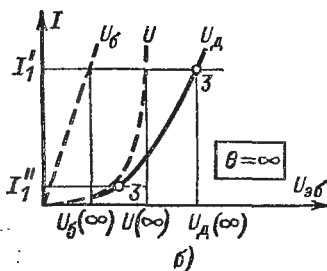
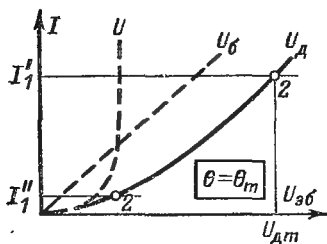
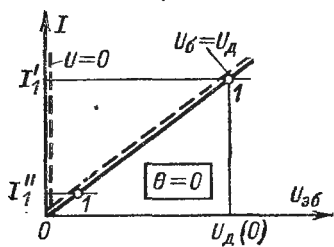
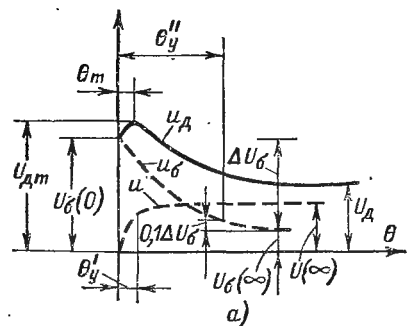


Рис. 2-41. Этап установления прямого напряжения.

а — временные диаграммы; б — динамические вольт-амперные характеристики диода (начальная, промежуточная и установившаяся) и соответствующие рабочие точки при малом и большом рабочих токах (пунктиром показаны составляющие полных характеристики — для перехода и для базового слоя).

зная закон модуляции сопротивления базы r_6 при высоком уровне инжекции. В § 2-8 было показано, что для достаточно толстой базы зависимость установившейся величины сопротивления r_6 от тока I_1 имеет вид (2-77а). При выводе этой формулы распределение носителей вдоль базы было принято экспоненциальным, как и должно быть в стационарном режиме. Во время переходного процесса распределение дырок характеризуется изображением, оригинал которого слишком сложен для последующего интегрирования. Поэтому примем, что распределение дырок все время остается экспоненциальным, но диффузионная длина L меняется со временем от значения $L = 0$ при $t = 0$ (или при $s = \infty$) до установившегося значения L при $t = \infty$ (или при $s = 0$). Такое допущение основано на выражениях (1-124) и (1-125) для операторной диффузионной длины и ее оригинала. Подставляя функцию $L(t)$ из (1-125) в выражение (2-77а), характеризующее толстую базу, и умножая обе части на ток I_1 , получаем искомое напряжение¹:

$$u_6(\theta) = U_6(0) \left[1 - \frac{L}{w} \ln(1 + \delta_1) \operatorname{erf} \sqrt{\theta} \right], \quad (2-96)$$

где $U_6(0) = I_1 r_{60}$ и $\delta_1 = I_1 / I_6$. Из формулы видно, что напряжение $u_6(\theta)$ монотонно уменьшается от начального значения $U_6(0)$ до установившегося значения $U_6(\infty)$, т. е. имеет н а ч а л ь н ы й в ы б р о с (рис. 2-41, а).

Значение выброса

$$\Delta U_6 = U_6(0) - U_6(\infty) = U_6(0) \frac{L}{w} \ln(1 + \delta_1).$$

Время установления можно оценить из условия $U_6(\theta_y) = U_6(0) - 0,9\Delta U_6$, которое дает:

$$\theta_y'' \approx 1,35.$$

Таким образом, длительность выброса несколько больше времени жизни τ .

Теперь остается просуммировать напряжения $u(\theta)$ и $u_6(\theta)$, чтобы получить полное прямое напряжение на диоде $u_d(\theta)$. В зависимости от тока I_1 функция $u_d(\theta)$ может быть разной. При небольшом токе, когда выброс ΔU_6 невелик, выброс напряжения u_d отсутствует. При достаточно большом токе кривая прямого напряжения имеет выброс, хорошо известный из практики диодных импульсных схем (рис. 2-41, а). Происхождение выброса дополнительно иллюстрируется на рис. 2-41, б.

Максимальное значение прямого напряжения U_{dm} при наличии выброса можно найти путем дифференцирования суммы функций $u(\theta)$ и $u_6(\theta)$, представленных выражениями (2-94) и (2-96). В результате оказывается (см. 3-е издание

¹ Вообще говоря, следовало бы подставить $L(\theta)$ не только вместо длины L , входящей в формулу (2-77а) непосредственно, но и вместо длины L , входящей в выражение для тока I_0 . Однако вторая подстановка практически несущественна.

данной книги), что время θ_m , соответствующее значению U_{dm} , близко к нулю. Соответственно

$$U_{dm} \approx U_6(0) = I_1 r_{60}; \quad (2-97)$$

$$\theta_y \approx \theta_y'' = 1,35. \quad (2-98)$$

Часто выброс прямого напряжения характеризуют не величиной U_{dm} , а производной величиной — *импульсным прямым сопротивлением* $R_{dm} = U_{dm}/I_1$, которое может в 1,5—3 раза превышать статическую величину R_d [см. (2-39)].

Рассасывание избыточных носителей. После переключения диода с прямого направления на обратное перераспределение зарядов в базе и переходе не может произойти мгновенно. Соответственно граничная концентрация носителей (2-13а) и связанное с ней

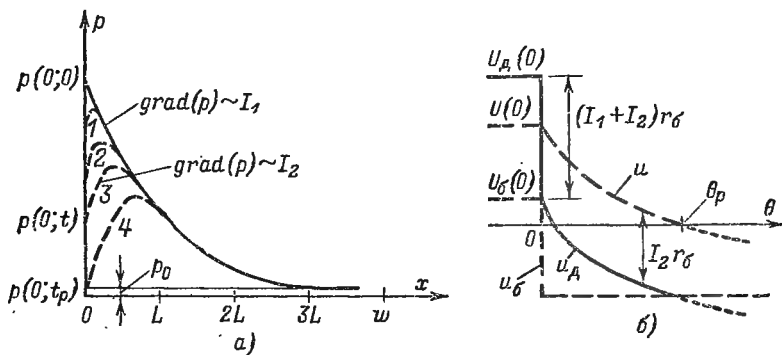


Рис. 2-42. Этап рассасывания при переключении.

а — распределение носителей в базе — исходное (сплошная линия) и для четырех последовательных моментов времени (пунктирные линии); б — временные диаграммы.

напряжение на переходе U уменьшаются постепенно (рис. 2-42, а), так что в течение некоторого времени на переходе сохраняется небольшое *п р я м о е* напряжение¹. Отсюда следует, что обратный ток в течение этого времени можно считать *з а д а н н о й* величиной, поскольку приложенное напряжение e_2 обычно много больше прямого напряжения на диоде:

$$I_2 = \frac{-e_2 - U_d}{R} \approx \frac{-e_2}{R} = \text{const.}$$

При анализе этапа рассасывания будем отсчитывать время от момента переключения. Примем, что при $t = 0$ избыточная концентрация дырок определяется стационарным распределением (2-31а).

¹ «Инерционность» прямого напряжения становится особенно наглядной, если представить себе переход зашунтированным барьерной и диффузионной емкостями, которые отражают перераспределение зарядов (см. начало данного параграфа) и которые, как и любой конденсатор, не могут перезаряжаться мгновенно.

С помощью (2-33) это распределение можно легко записать следующим образом:

$$\Delta p(x; 0) = p_0 \frac{I_1}{I_0} e^{-\frac{x}{L}}. \quad (2-99)$$

Изображением производной $\frac{\partial(\Delta p)}{\partial t}$ будет $s[\Delta p - \Delta p(x; 0)]$, и диффузионное уравнение (1-79а) в операторной форме приводится к виду

$$\frac{d^2(\Delta p)}{dx^2} - \frac{1+s\tau}{L^2} \Delta p = -\frac{s}{D_p} \Delta p(x; 0). \quad (2-100)$$

Частным решением уравнения является функция $\Delta p = \Delta p(x; 0)$; это легко проверить подстановкой выражения (2-99). Общее решение имеет прежний вид (2-90). Исходя из граничных условий

$$\begin{aligned} -SqD_p \left| \frac{\partial(\Delta p)}{\partial x} \right|_{x=0} &= -I_2; \\ \Delta p(\infty) &= 0, \end{aligned}$$

после преобразований получим изображение избыточной концентрации:

$$\Delta p(x; s) = p_0 \left(\frac{I_1}{I_0} e^{-\frac{x}{L}} - \frac{I_1 + I_2}{I_0 \sqrt{1+s\tau}} e^{-\frac{x}{L} \sqrt{1+s\tau}} \right). \quad (2-101)$$

По таблицам находим оригинал этой функции для $x = 0$:

$$\Delta p(0; \theta) = p_0 \left(\frac{I_1}{I_0} - \frac{I_1 + I_2}{I_0} \operatorname{erf} \sqrt{\theta} \right). \quad (2-102)$$

Приравняв (2-102) и (2-28а), легко найти напряжение на переходе (рис. 2-42, б):

$$u(\theta) = \varphi_T \ln \left(1 + \frac{I_1}{I_0} - \frac{I_1 + I_2}{I_0} \operatorname{erf} \sqrt{\theta} \right). \quad (2-103)$$

Время рассасывания θ_p обычно определяют, полагая $U(\theta_p) = 0$, т. е. считая, что этап рассасывания заканчивается тогда, когда граничная концентрация носителей достигает равновесного значения. При этом время θ_p получается в неявном виде:

$$\operatorname{erf} \sqrt{\theta_p} = \frac{I_1}{I_1 + I_2}. \quad (2-104а)$$

Используя аппроксимацию (2-93а), можно получить время θ_p в явном виде:

$$\theta_p = -\ln \left[1 - \left(\frac{I_1}{I_1 + I_2} \right)^2 \right]. \quad (2-104б)$$

Например, если $I_2 = I_1$, то $\theta_p \approx 0,3$. С увеличением запирающего тока I_2 время рассасывания уменьшается, что вполне естественно.

Полное напряжение на диоде в момент θ_p равно $-I_2 r_6$ (рис. 2-42, б). Значение r_6 в первом приближении можно рассчитать по формуле (2-77а), поскольку за время θ_p заряд дырок в базе меняется незначительно (см. рис. 2-42, а). До переключения на базе падало напряжение $+I_1 r_6$. Поэтому в момент переключения диода получается так называемый «омический» скачок напряжения:

$$\Delta U_d = \Delta U_6 = -(I_1 + I_2) r_6.$$

Теперь рассмотрим случай, когда $I_2 = 0$, т. е. когда диод не переключается, а отключается (см. рис. 2-40, правая часть). В момент отключения «омический» скачок напряжения равен $-I_1 r_6$, после чего напряжение U_6 остается равным нулю,

Рис. 2-43. Этап рассасывания при выключении.

так как $I_2 = 0$. Следовательно, после отключения имеем $U_d = U$ и напряжение на переходе согласно (2-103) меняется следующим образом (рис. 2-43):

$$u(\theta) = \varphi_T \ln \left[1 + \frac{I_1}{I_0} (1 - \operatorname{erf} \sqrt{\theta}) \right]. \quad (2-105)$$

Если в квадратных скобках выделить множитель $1 + I_1/I_0$ и прологарифмировать полученное произведение, то выражение (2-105) можно привести к виду

$$u(\theta) = U(0) + \varphi_T \ln(1 - \operatorname{erf} \sqrt{\theta}), \quad (2-106a)$$

где для простоты множитель $I_1/(I_0 + I_1)$ при $\operatorname{erf} \sqrt{\theta}$ положен равным единице. Применяя аппроксимацию (2-93а), получаем переходную характеристику в форме, удобной для построения и анализа:

$$u(\theta) = U(0) + \varphi_T \ln(1 - \sqrt{1 - e^{-\theta}}). \quad (2-106б)$$

В начале переходного процесса, когда $\theta < 0,5$, можно воспользоваться аппроксимацией (2-93б) и разложить логарифм в ряд, ограничиваясь членом первого порядка; тогда

$$u(\theta) \approx U(0) - \varphi_T \sqrt{\theta}. \quad (2-107a)$$

При достаточно больших значениях θ используем аппроксимацию (2-93в) и получаем линейную характеристику:

$$u(\theta) \approx U(0) - \varphi_T (\theta + \ln 2). \quad (2-107б)$$

В самом конце процесса, когда второе слагаемое в квадратных скобках выражения (2-105) меньше единицы, разлагая логарифм в ряд и используя после этого аппроксимацию (2-93в), получаем экспоненциальный участок характеристики:

$$u(\theta) \approx \frac{\varphi_T}{2} \cdot \frac{I_1}{I_0} e^{-\theta}. \quad (2-107в)$$

На этом участке напряжение U не превышает долей φ_T , т. е. пренебрежимо мало. Поэтому основным участком характеристики является линейный.

Из формулы (2-107б) можно оценить время рассасывания, полагая $U(\theta_p) = 0$ и пренебрегая слагаемым $\ln 2$ (это дает некоторый «запас», учитывающий часть экспоненциального «хвоста»):

$$\theta_p = \frac{U(0)}{\Phi_T} \approx \ln \frac{I_1}{I_0}. \quad (2-108)$$

Например, если $I_1/I_0 = 10^5$, то $\theta_p \approx 11$. Как видим, время рассасывания в случае выключения несравненно больше, чем в случае переключения (даже если последнее осуществляется сравнительно небольшими обратными токами I_2).

Восстановление обратного тока (сопротивления). Этот участок переходной характеристики (рис. 2-44) наиболее сложен для анализа. Строгое решение [46] дает три разных закона изменения тока в трех смежных интервалах времени, что, конечно, крайне неудобно на практике.

Приводимое ниже решение исходит из вполне оправданного допущения о том, что в начале рассматриваемого участка (при $t = 0$) распределение дырок в базе выражается разностью двух экспонент¹:

$$\Delta p(x, 0) = p_0 \frac{I_1}{I_0} \left(e^{-\frac{x}{L}} - e^{-\frac{x}{l}} \right), \quad (2-109)$$

где $l < L$.

Функция (2-109) равна нулю при $x = 0$ и при $x = \infty$ и имеет максимум между $x = 0$ и $x = L$, что соответствует физической картине процессов в толстой базе в конце стадии рассасывания (см. кривую d на рис. 2-42, а). Эквивалентную длину l определим, исходя из граничного условия

$$-SqD_p \left| \frac{\partial(\Delta p)}{\partial x} \right|_{x=0} = -I_2,$$

куда подставим производную от концентрации (2-109). Вычисления приводят к следующему результату: $l = L \frac{I_1}{I_1 + I_2}$.

Из соотношения $l < L$ следует, что вторая экспонента в выражении (2-109) уменьшается с ростом x быстрее, чем первая; значит, вдали от перехода распределение дырок близко к тому, которое было при прямом токе I_1 до переключения. Это также представляется физически справедливым.

Опуская несколько громоздкие выкладки, связанные с решением диффузионного уравнения (2-100), заметим лишь, что уда-

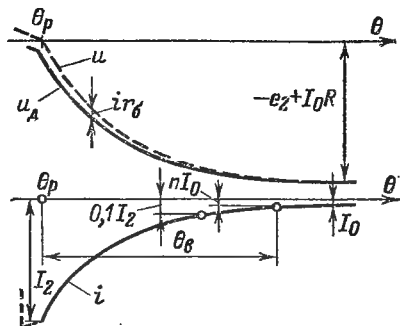


Рис. 2-44. Этап восстановления обратного тока и обратного напряжения.

¹ Такая аппроксимация при $x = 0$ дает $\Delta p = 0$ вместо $\Delta p = -p_0$. Эта погрешность практически не сказывается на результатах.

ется получить изображение $U(s)$ в следующей форме¹:

$$U(s) = I_0 R \left(\sqrt{1 + s\tau} - \frac{I_2}{I_0} \frac{B+1}{B + \sqrt{1 + s\tau}} \right), \quad (2-110)$$

где

$$B = \frac{L}{I} = 1 + \frac{I_2}{I_1}.$$

По таблицам находим оригинал $u(\theta)$, который выражает обратное напряжение в виде е д и н о й функции для всего рассматриваемого участка:

$$u(\theta) = I_0 R \left[\frac{e^{-\theta}}{\sqrt{\pi\theta}} + \left(1 + \frac{I_1}{I_0} \right) \operatorname{erf} \sqrt{\theta} - \right. \\ \left. - \frac{I_1 + I_2}{I_0} (1 - e^{(B^2 - 1)\theta}) \operatorname{cerf} \sqrt{B^2\theta} \right]. \quad (2-111)$$

Первый член в квадратных скобках существен лишь в самом начале кривой, обычно при $\theta < 0,01$; функции erf и $\operatorname{cerf} = 1 - \operatorname{erf}$ можно аппроксимировать с помощью (2-93). Таким образом, функция $u(\theta)$ практически не очень сложна. Сравнение с существующими строгими решениями показывает хорошее совпадение.

Учитывая, что на участке восстановления обратный ток изменяется по закону

$$i(\theta) = -I_2 - \frac{u(\theta)}{R},$$

не представляет труда записать переходную характеристику обратного тока:

$$i(\theta) = -I_0 \left[\frac{e^{-\theta}}{\sqrt{\pi\theta}} + \left(1 + \frac{I_1}{I_0} \right) \operatorname{erf} \sqrt{\theta} - \right. \\ \left. - \frac{I_1}{I_0} + \frac{I_1 + I_2}{I_0} e^{(B^2 - 1)\theta} \operatorname{cerf} \sqrt{B^2\theta} \right]. \quad (2-112)$$

Время восстановления обратного тока θ_v нельзя определять на уровне $0,1I_2$, так как величина $0,1I_2$ может немного превышать I_0 , что не соответствует понятию запертого состояния диода (рис. 2-44). Поэтому определим время θ_v , задаваясь определенным значением обратного тока $i(\theta_v) = -nI_0$. Пренебрежем в квадратных скобках выражения (2-112) первым членом и используем аппроксимацию (2-93в). Тогда после преобразований можно получить время θ_v в следующей простой явной форме²:

$$\theta_v = \ln \frac{I_2}{2(n-1)I_0}. \quad (2-113)$$

¹ Путь решения указан в [45], а также во 2-м издании данной книги. В качестве граничных условий приняты: $\Delta p(0) = -p_0$ и $\Delta p(\infty) = 0$.

² В данном случае аппроксимация (2-93а) приводит к завышению времени восстановления в 1,5—2 раза. Более строгая аппроксимация приводит к неудобному трансцендентному решению.

Например, если $I_2/I_0 = 10^5$ и $n = 3$, то $\theta_v = 10$. Практически при $\tau = 1$ мкс время восстановления может составлять 5—7 мкс. У специальных импульсных диодов время жизни и время восстановления в 50—200 раз меньше.

Особый интерес представляют диоды с накоплением заряда (ДНЗ) [44], у которых время восстановления значительно меньше времени рассасывания, так что выброс обратного тока имеет почти прямоугольную форму. Такое соотношение времен t_v и t_p (обратное по отношению к обычным диодам) достигается благодаря наличию тормозящего электрического поля в базе. Тормозящее поле способствует скоплению инжектированных носителей вблизи эмиттера и тем самым — малому остаточному заряду в конце этапа рассасывания. Внутреннее тормозящее поле обеспечивается неоднородностью базы. При достаточно большом обратном токе ($I_2/I_1 > 10$) отношение t_v/t_p у ДНЗ может составлять 0,2—0,3 и менее. Такая прямоугольная форма импульсов используется в генераторах гармоник, умножителях частоты, диодных усилителях, формирователях импульсов и других схемах.

Глава третья

РАЗНОВИДНОСТИ ПОЛУПРОВОДНИКОВЫХ ДИОДОВ

Теория и свойства плоскостных полупроводниковых диодов, изложенные в предыдущей главе, лежат в основе всех других типов диодов, количество которых в настоящее время довольно велико. Специфика каждого из этих специальных диодов требует особого анализа, но мы ограничимся их качественной характеристикой в той мере, в какой это полезно при разработке транзисторов и интегральных схем.

3-1. ТОЧЕЧНЫЕ ДИОДЫ

Точечные (или точечно-контактные) диоды (рис. 3-1) появились намного раньше плоскостных. Однако процессы, происходящие в них, сложнее и до сих пор полностью не изучены.

Эффект выпрямления на границе между металлической иглой и пластинкой полупроводника имеет место даже при простом их соприкосновении. Однако более надежный точечный контакт получают путем так называемой *формовки* [47]. Процесс формовки заключается в пропускании через собранный точечный диод сравнительно мощных, но коротких импульсов тока в прямом или обратном направлении. Количество, величина и длительность формирующих импульсов выбирают на основании опытных данных. Общей целью формовки является сильный местный нагрев контакта, при котором происходит своеобразное сплавление кончика иглы с полу-

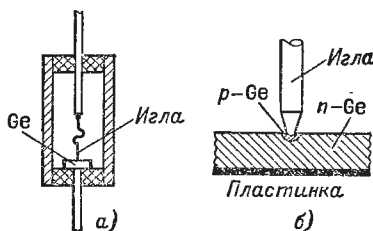


Рис. 3-1. Конструкция точечного диода (а) и структура его перехода (б).

проводником. Сплавление обеспечивает стабильность и механическую прочность контакта, что и было первоначальной целью формовки. Однако, как выяснилось позднее, при сплавлении происходит также изменение типа проводимости в тонком слое полупроводника, прилегающем к игле. Это превращение объясняется диффузией определенных примесей из иглы в полупроводник при сильном разогреве и частичном расплавлении обоих элементов в месте контакта. Например, в случае, когда исходная пластинка германия имеет электронную проводимость, а материалом иглы является бериллиевая бронза, может происходить диффузия бериллия в германий. Бериллий, будучи акцептором по отношению к германию, обуславливает наличие тонкого p -слоя в германии вблизи иглы (рис. 3-1). Как видим, в данном случае получается p - n переход, правда, своеобразной (не плоской) конфигурации и малой площади. Обычно при анализе форму перехода в точечном диоде принято считать полусферической, что, несомненно, близко к действительности¹. Отметим характерные особенности точечных переходов и диодов.

Очевидно, что малая площадь перехода обуславливает малую емкость перехода, но и малую допустимую мощность. Значения допустимых прямых токов у точечных диодов значительно меньше, чем у плоскостных, а эффективное сопротивление базы больше из-за малой площади эмиттера². Сопротивления базы у точечных диодов составляют десятки и сотни ом, допустимые мощности рассеяния около 10 мВт, а прямые токи обычно не больше 10—20 мА. Существенное превышение допустимого тока (даже в течение короткого интервала времени) приводит у точечных диодов к «переформовке» контакта и ухудшению или утрате выпрямляющих свойств.

Для прямой вольт-амперной характеристики точечного диода выражение (2-33) недействительно, так как даже при очень небольших токах уровень инжекции (2-69) оказывается весьма высоким из-за малой площади перехода. Для начального участка характеристики приемлемо выражение (2-79б), в котором $m \approx 2$. При больших

¹ Возникновение p - n перехода в процессе формовки — весьма наглядный, но не единственный механизм, объясняющий вентильные свойства точечного диода. В общем случае формовка обеспечивает лишь хороший контакт иглы с пластинкой, а нелинейность вольт-амперной характеристики обусловлена наличием обедненного приповерхностного слоя в системе «металл-полупроводник» (см. рис. 2-15). Для этого, если задан тип полупроводника (кремний, германий и т. п.), необходимо, разумеется, подбирать металл с соответствующей работой выхода. Еще один вариант точечного диода (не связанный с формовкой) основан на образовании обедненного слоя благодаря достаточно высокой плотности поверхностных состояний (см. с. 127). В этом случае игла играет роль простого омического контакта с поверхностью полупроводника.

² Сопротивление базы при полусферическом эмиттере выражается следующим образом:

$$r_b = \frac{\rho_b}{\pi d},$$

где d — диаметр эмиттерной полусферы (обычно около 10—20 мкм).

тока уровня инжекции получается настолько высоким, что экспоненциальная характеристика вырождается и приближается к параболической (2-80). Такому вырождению способствует сравнительно большое отношение ρ_a/ρ_b , так как эмиттер специально не легирован.

Особенно сильно различаются статические характеристики точечных и плоскостных диодов в области обратных токов (рис. 3-2). Поскольку площадь перехода мала, мал и тепловой ток I_0 . Однако участок насыщения небольшой и обычно плохо заметен, так как уже при сравнительно небольших напряжениях (несколько вольт) обратный ток существенно возрастает за счет утечек, а также за счет заметного повышения температуры перехода (теплоотвод затруднен малой площадью контакта, и тепловое сопротивление R_t велико).

Важной особенностью обратной характеристики является участок с отрицательным дифференциальным сопротивлением, который обусловлен тепловым пробоем (см. § 2-7, рис. 2-30). Этот участок по праву называют «участком Лосева», так как советский физик О. В. Лосев [48] первым использовал его в 1922 г. в схемах усилителей и генераторов (кристадины). Несмотря на принципиальную ценность указанного участка, его практическое использование нецелесообразно вследствие большого разброса значения отрицательного сопротивления и координат начала и конца участка, плохой стабильности во времени и малого срока службы диода в таком перенапряженном режиме. Кроме того, частотный предел в схемах, использующих участок Лосева, очень невелик (десятки — сотни килогерц) из-за инерционности тех тепловых процессов, которые лежат в основе отрицательного сопротивления. Однако основная идея О. В. Лосева — возможность усиления и генерации с помощью полупроводниковых двухполосников — широко реализуется с помощью новых типов диодов, например диодов Ганна [8] и туннельных (см. § 3-3).

Переходные процессы протекают в точечных диодах качественно так же, как и в плоскостных (см. § 2-9). Количественные различия связаны с меньшими площадью перехода и временем жизни. Время жизни у точечных диодов меньше, чем у плоскостных, так как роль поверхностной рекомбинации возрастает с уменьшением площади перехода. Поскольку длительность всех этапов переключения пропорциональна времени жизни (см. § 2-9, где $\theta = t/\tau$), при прочих равных условиях повышается скорость переходных процессов и расширяется рабочий диапазон частот. Поэтому точечные диоды до сих пор используются в высокочастотных схемах. Однако за последние несколько лет в связи с освоением технологии диодов Шоттки (см. § 3-4) и тенденцией использовать групповые

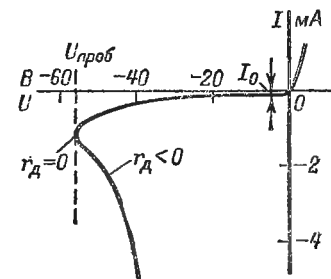


Рис. 3-2. Статическая характеристика точечного диода.

методы изготовления, свойственные микроэлектронике, точечные диоды утрачивают свое значение в электронной технике.

Обширный и своеобразный класс точечных полупроводниковых диодов составляют германиевые и кремниевые детекторные диоды, применяющиеся главным образом в технике СВЧ [16, 49]. Технологическая особенность детекторных диодов заключается в весьма низкоомном материале базы и очень малом (2—3 мкм) радиусе точечного контакта. Низкое удельное сопротивление базы (примерно 0,01—0,001 Ом·см) обуславливает очень малое время жизни носителей, что вместе с малой площадью контакта объясняет хорошие частотные свойства, необходимые для работы в области СВЧ.

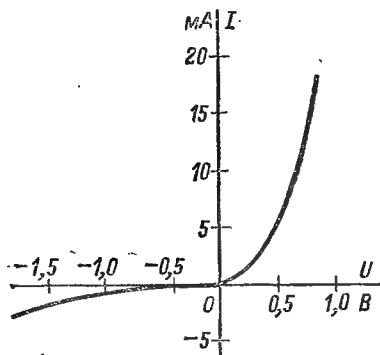


Рис. 3-3. Статическая характеристика детекторного диода.

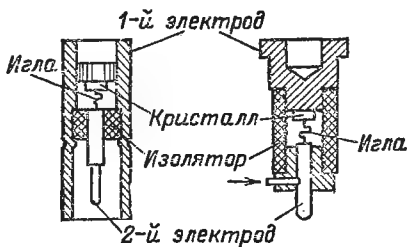


Рис. 3-4. Конструктивные варианты детекторных диодов.

Однако по параметрам вольт-амперной характеристики детекторные диоды значительно уступают высоковольтным точечным диодам (рис. 3-3). Напряжение пробоя у детекторных диодов составляет всего 3—5 В, так как при малом удельном сопротивлении базы переход получается очень узким и напряженность поля в нем велика даже при малых напряжениях. Тепловой ток, пропорциональный величине ρ_b (см. § 2-6), весьма мал, но рост обратного тока начинается практически с нуля, так как при очень узком переходе носители проникают через потенциальный барьер вследствие туннельного эффекта. Прямые напряжения из-за малого значения I_0 относительно велики [см. (2-37)].

Отмеченные недостатки не столь существенны для детекторных диодов, так как смесители и преобразователи СВЧ, в которых они используются, работают от малых сигналов (сотые доли вольта¹ и меньше). Для таких диодов помимо рабочей частоты важны такие параметры, как коэффициент шума, коэффициент преобразования и др. Эти параметры могут быть достаточно хорошими, несмотря на «плохую» (с точки зрения «обычных» диодов и обычных электронных схем) вольт-амперную характеристику.

¹ Обычно в области СВЧ удобнее не «расчленять» уровень сигнала на напряжение и ток, а оперировать мощностью.

Конструкция детекторных диодов обычно приспособлена к соединению с элементами волноводов, измерительных головок и других деталей систем СВЧ (рис. 3-4).

3-2. ПОЛУПРОВОДНИКОВЫЕ СТАБИЛИТРОНЫ

Аналогом газонаполненного стабилизатора напряжения (стабилитрона) служит кремниевый плоскостной диод, работающий в области пробоя [16]. Такой «полупроводниковый стабилитрон» по большинству параметров превосходит газоразрядный (табл. 3-1) и, главное, может иметь несравненно меньшие рабочие напряжения, что совершенно необходимо в транзисторных схемах, практически всегда низковольтных.

Выбор кремния в качестве материала для полупроводниковых стабилитронов обусловлен главным образом малым обратным током. При этом саморазогрев диода в предпробойной области отсутствует и переход в последнюю получается достаточно резким (рис. 3-5). Кроме того, в самой области пробоя, даже при большом токе, нагрев диода не носит лавинообразного характера, так как токи I_0 и I_G остаются малыми в широком температурном диапазоне [роль этих токов в тепловом пробое была подчеркнута в связи с формулой (2-60)]. Соответственно рабочий участок характеристики идет почти вертикально и не имеет отрицательного наклона, характерного для теплового пробоя (см. рис. 2-30).

Механизм пробоя в полупроводниковых стабилитронах может быть туннельным, лавинным или смешанным в зависимости от удельного сопротивления базы (см. § 2-7). У стабилитронов с весьма низкоомной базой (низковольтных) имеет место туннельный (зенеровский) пробой, а у стабилитронов с высокоомной базой (сравнительно высоковольтных) — лавинный пробой.

Заметим, что в случае лавинного пробоя выгодно делать базу из электронного полупроводника, так как при этом излом обратной характеристики будет более резким [см. (2-55)] и значения показателя n в табл. 2-2].

Большой интерес представляет температурная зависимость стабилизированного напряжения в виде температурной чувствительности $\varepsilon = dU_{\text{проб}}/dT$ или температурного коэффициента напряжения $\text{ТКН} = \varepsilon/U_{\text{проб}}$. У диодов с туннельным пробоем (низковольтных) пробивное напряжение находится в прямой зависимости

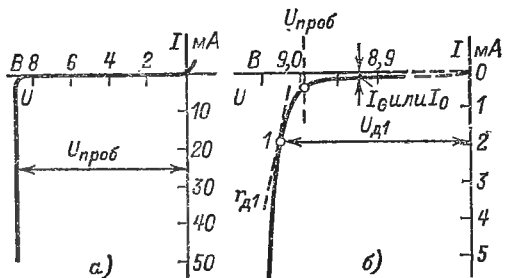


Рис. 3-5. Статическая характеристика полупроводникового стабилитрона.

а — полная характеристика; б — рабочий участок.

Наименование прибора	$U_{\text{раб}}, \text{ В}$	$\delta U_{\text{раб}}, \%$
Газонаполненные стабилитроны	70—150	± 5
Прецизионные полупроводниковые стабилитроны	6—20	± 5
Полупроводниковые стабилитроны малой мощности	2—200	± 5
Полупроводниковые стабилитроны средней мощности	3—200	± 5
Полупроводниковые стабилитроны большой мощности	5—200	± 10

Примечание. $U_{\text{раб}}$ — рабочее напряжение; $\delta U_{\text{раб}}$ — допуск на номинал рабочего напряжения при токе, указанном в скобках; $R_{\text{обр}}$ — обратное сопротивление; $P_{\text{расс}}$ — мощность рассеяния.

от ширины запрещенной зоны. Величина φ_z , как отмечалось на стр. 26, уменьшается с повышением температуры (температурный коэффициент напряжения около 0,01 %/град), а значит, уменьшается и пробивное напряжение согласно (2-51а). Следовательно, температурная чувствительность в этом случае отрицательна: $\varepsilon < 0$.

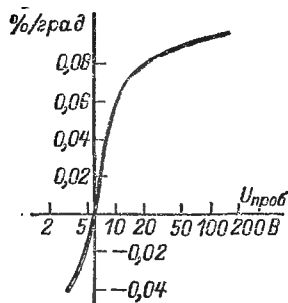


Рис. 3-6. Зависимость ТКН на рабочем участке кремниевого стабилитрона от напряжения пробоя.

У диодов с лавинным пробоем (сравнительно высоковольтных) пробивное напряжение согласно (2-56) находится в обратной зависимости от подвижности носителей, поскольку $\rho \sim \mu^{-1}$. Учитывая (1-32), приходим к выводу, что в этом случае пробивное напряжение увеличивается с ростом температуры; соответственно температурная чувствительность будет положительной: $\varepsilon > 0$.

Зависимость ТКН от $U_{\text{проб}}$ показана на рис. 3-6, откуда следует, что граничным напряжением, при котором $\varepsilon \approx 0$, является $U_{\text{проб}} \approx 5,5 \text{ В}$. Большинство стабилитронов имеет более высокие рабочие напряжения, поэтому для них типичны положительные ε порядка нескольких милливольт на градус. Это значение при широком диапазоне изменений температуры (50—100° С и более) часто оказывается неприемлемым. Тогда применяют последовательное соединение полупроводниковых стабилитронов с диодами, работающими в прямом направлении, в котором температурная чувствительность обычно отрицательна (см. § 2-8). Такое решение используется в прецизионных стабилитронах (например, в диодах Д818 и КС196).

Величина ε практически не зависит от тока, если последний превышает 1—2 мА. При меньших токах кремниевые стабилитроны

полупроводниковых стабилитронов

$I_{\text{макс}}$ (до +50 °С), мА	$R_{\text{внутр}}$, Ом	Температурный коэффициент напряжения, %/град	$R_{\text{обр}}$, МОм	$T_{\text{доп}}$, °С	$P_{\text{доп}}$ (до 50 °С), Вт
30—40	75—150 ($I_{\text{макс}}$)	$\pm 0,01 \div \pm 0,07$	5	+70	8
10—20	5—15 (10 мА)	$0,0002 \div 0,01$	—	+100	0,1—0,2
20—50	10—50 (10 мА)	$\pm 0,02 \div \pm 0,08$	2—20	+150	0,3—0,5
25—500	2—200 ($I_{\text{макс}}$)	$0,05 \div 0,12$	2—100	+150	1—5
250—10 ⁴	2—60 ($I_{\text{макс}}$)	$0,07 \div 0,15$	—	—	10—50

чего напряжения; $I_{\text{макс}}$ — максимальный рабочий ток; $R_{\text{внутр}}$ — внутреннее дифференциальное сопротивление; $T_{\text{доп}}$ — допустимая окружающая температура; $P_{\text{доп}}$ — допустимая мощность рассеяния.

не рекомендуется применять, главным образом из-за роста дифференциального сопротивления диода r_d на пробивном участке. При токе 1 мА сопротивление увеличивается в среднем в 2 раза (а иногда и больше) по сравнению с номинальным значением, что нежелательно в схемах стабилизаторов (см. гл. 23). Кроме того, при малых токах резко возрастают собственные шумы диода.

При увеличении тока вплоть до максимального значения¹ сопротивление уменьшается, но весьма незначительно (на 10—20%). Зависимость $r_d(I)$ в области пробоя объясняется неоднородностью p - n перехода: отдельные участки его пробиваются поочередно по мере повышения напряжения [42]. Удельное сопротивление «пробитых» участков, близкое к ρ_b , на много порядков меньше, чем удельное сопротивление «непробитых» (запертых) участков, и поэтому определяет значение r_d . С увеличением площади «пробитых» участков сопротивление r_d уменьшается и в пределе стремится к значению r_b , которое достигается примерно при номинальном токе.

Временная нестабильность опорного напряжения весьма мала. В первые 5—10 мин после включения она не превышает 0,2—0,5%, а в дальнейшем (при постоянной температуре) лежит в пределах до 1 мВ, т. е. составляет менее 0,01%. Такая высокая стабильность

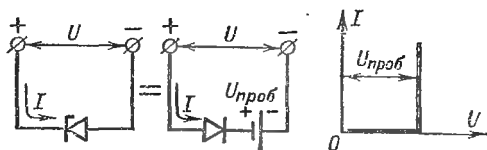


Рис. 3-7. Эквивалентная схема полупроводникового стабилитрона в области пробоя.

¹ Максимальный ток полупроводникового стабилитрона ограничен допустимой мощностью рассеяния. С ростом температуры допустимая мощность, а значит, и допустимый ток уменьшаются (например, для диодов Д808—Д813 допустимая мощность уменьшается с 300 мВт при +50°С до 70 мВт при +120°С).

позволяет во многих случаях использовать полупроводниковый стабилитрон в качестве нормального элемента.

Поскольку пробойный режим не связан с инжекцией неосновных носителей, в полупроводниковом стабилитроне отсутствуют инерционные явления (накопление и рассасывание носителей) при переходе из области пробоя в область запира- ния и обратно. Эта особенность делает возможным применение полупроводниковых стабилитронов не только в стабилизаторах напряжения, но и в импульсных схе- мах. К числу таких схем относятся ограничители и фиксаторы уровня. В них полупроводниковый стабилитрон совмещает функции источника фиксирующего напряжения $U_{\text{проб}}$ и обычного диода с весьма резким изломом характеристики (рис. 3-7). Кроме того, полупроводниковые стабилитроны могут применяться в качестве шунтов, защищающих от перенапряжений, в качестве элементов меж- каскадной связи в усилителях постоянного тока, триггерах и других схемах [50].

3.3. ТУННЕЛЬНЫЕ ДИОДЫ

Туннельные диоды [16, 49, 51, 52], разработанные в 1958— 1959 гг. по предложению японского физика Есаки, интересны тем, что будучи двухполюсниками, они могут усиливать сигналы по- добно транзисторам. Это объясняется наличием участка с отрица- тельным сопротивлением на их вольт-амперной характеристике.

В отличие от «участка Лосева» у точечных диодов (см. § 3-1) отрица- тельное сопротивление у туннельных диодов имеется не на обратной, а на прямой ветви характеристики (см. рис. 3-8; отрицательное сопротивление ме- жду точками 1 и 2). Рассмотрим проис- хождение такой формы характери- стики, воспользовавшись энергетическими диаграммами p - n перехода.

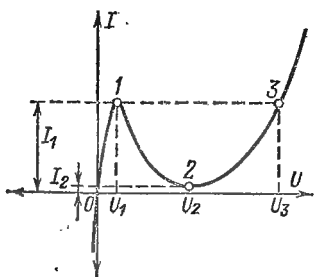


Рис. 3-8. Статическая харак- теристика туннельного дио- да.

Отличительными особенностями тун- нельного диода являются очень малые удельные сопротивления p - и n -слоев и соответственно очень малая ширина пе- рехода. Концентрация примесей в слоях достигает 10^{19} см^{-3} и больше.

В этом случае полупроводник вырож- дается, превращаясь в полуметалл (см. с. 20). Уровни примесных атомов сливаются в зоны, а последние в свою очередь сливаются с соответствующими основными зонами слоев. В результате уровни Ферми, как и в металле, располагаются не в запрещенных зонах p - и n -слоев, а в разрешенных зонах: в валентной зоне p -слоя и в зоне проводимости n -слоя. При этом энергетическая диаграмма с и м - м е т р и ч н о г о перехода в равновесном состоянии будет примерно такой, как показано на рис. 3-9, а. Как видим, нижняя часть зоны проводимости в слое n и верхняя часть валентной зоны в слое p оказались разделенными весьма узким запиорным слоем. Если ши- рина его не превышает 0,01—0,02 мкм (что легко получить в низ- коомном переходе), то носители имеют возможность переходить

в смежный слой «по горизонтали», т. е. не преодолевая потенциального барьера. Это явление обусловлено туннельным эффектом (§ 2-7), откуда и происходит название диодов.

Ниже мы рассмотрим работу диода, анализируя движение электронов. Поведение дырок совершенно аналогично, а относительная роль обоих типов носителей, как обычно, зависит от степени симметрии диода, т. е. от соотношения

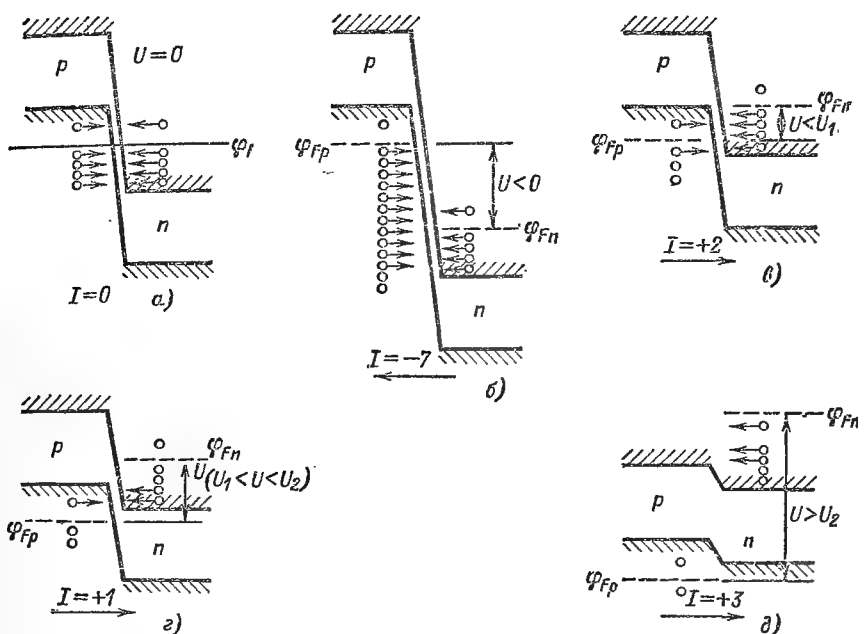


Рис. 3-9. Энергетические диаграммы туннельного диода на разных участках характеристики.

a — равновесное состояние; *б* — обратное включение; *в* — прямое включение при $U < U_1$; *г* — прямое включение при $U_1 < U < U_2$; *д* — прямое включение при $U > U_2$.

удельных сопротивлений слоев. Распределение электронов по энергиям отражено на рис. 3-9 разными расстояниями между кружками, обозначающими электроны. Стрелками снабжены те электроны, которые способны перейти в смежный слой тем или иным путем. Результирующий ток через переход оценивается на рис. 3-9 как разность электронных потоков, проходящих из одного слоя в другой.

В равновесном состоянии потоки электронов в обоих направлениях уравниваются и ток отсутствует (рис. 3-9, *a*).

Приложим к диоду внешнее напряжение обратной полярности (т. е. плюсом к *n*-слою). Энергетическая диаграмма для этого случая показана на рис. 3-9, *б*. Так как количество электронов с энергией, превышающей уровень Ферми, невелико, то поток электронов из *p*-слоя в *n*-слой увеличится, а обратный поток останется почти

изменением потока электронов, результирующий ток будет протекать в направлении от n -слоя к p -слою. Этот ток очень быстро возрастает с увеличением обратного напряжения (см. рис. 3-8), поскольку плотность электронов в глубине валентной зоны огромна и малейшее приращение разности $\varphi_{F_p} - \varphi_{F_n}$ сопровождается существенным изменением потока электронов из p -слоя в n -слой.

Теперь приложим к диоду небольшое прямое напряжение. Энергетическая диаграмма для этого случая показана на рис. 3-9, в. Легко заметить, что поток электронов из p -слоя в n -слой сильно убывает, а обратный поток меняется сравнительно слабо. Следовательно, результирующий ток протекает в направлении от p -слоя к n -слою и при небольших прямых напряжениях возрастает с увеличением напряжения (рис. 3-8). Граница этого участка приблизительно соответствует диаграмме на рис. 3-9, в, на которой уровень φ_{F_n} совпадает с потолком валентной зоны p -слоя.

При дальнейшем увеличении прямого напряжения поток электронов из n -слоя в p -слой убывает (рис. 3-9, г) и соответственно убывает прямой ток. В результате на вольт-амперной характеристике получается участок с отрицательным сопротивлением (рис. 3-8). Конец этого участка соответствует такому напряжению, при котором потолок валентной зоны в p -слое совпадает с дном зоны проводимости в n -слое. При еще большем напряжении запрещенная зона делается «сквозной», туннельный эффект исчезает и ток снова увеличивается, но уже за счет обычного механизма — преодоления электронами потенциального барьера (рис. 3-9, д).

Таким образом, вольт-амперная характеристика туннельного диода (рис. 3-8) складывается из двух частей: *туннельной* (левее точки 2) и *диффузионной* (правее точки 2). Последняя, как обычно, обусловлена инжекцией и описывается формулой (2-33). Однако напряжения на диффузионной ветви туннельного диода получаются заметно больше, чем у обычного диода из того же материала. Это объясняется тем, что равновесная высота потенциального барьера у туннельных диодов больше примерно на 2 ($\varphi_{сд} + \varphi_{ва}$), где $\varphi_{сд}$ и $\varphi_{ва}$ — уровни доноров и акцепторов в не вырожденном полупроводнике, отсчитанные соответственно от дна зоны проводимости и потолка валентной зоны. Следует также отметить, что если бы работа туннельного диода строго соответствовала зонным диаграммам на рис. 3-9, то минимальный ток в точке 2 был бы намного меньше, чем это имеет место в действительности. Происхождение «избыточного» минимального тока приписывают наличию ловушек в запрещенной зоне. Такие ловушки могут обеспечить туннельный эффект даже после образования «сквозной» запрещенной зоны и тем самым затягивают спадающий туннельный участок характеристики [53].

Статические свойства туннельного диода удобно характеризовать координатами экстремальных точек (1 и 2 на рис. 3-8), а также напряжением в точке 3, соответствующим току I_1 .

Зонные диаграммы показывают, что пиковое напряжение U_1 должно находиться в прямой зависимости от глубины залегания уровней Ферми в слоях. Поэтому оно больше у арсенида галлия, у которого уровни примесей расположены дальше от краев запрещенной зоны, чем у германия¹.

Вместе с напряжением U_1 возрастает пиковый ток I_1 , так как крутизна восходящего участка туннельной ветви меняется сравнительно мало. Напряжение «провала» U_2 и «диффузионное» напряжение U_3 возрастают с увеличением концентрации примесей и ширины запрещенной зоны. Поэтому арсенидо-галлиевые диоды характерны большим диапазоном токов и напряжений, чем германиевые.

В зависимости от схемы применения желательное значение пикового тока может быть разным. Что касается минимального тока I_2 , то он всегда должен быть по возможности малым, с тем чтобы отношение I_1/I_2 было большим и участок с отрицательным сопротивлением был ярко выражен.

Изменения температуры влияют на статическую характеристику и ее параметры сравнительно слабо, так как оба слоя являются вырожденными полупроводниками — полуметаллами. Однако такое влияние все же имеется, поскольку положение уровней Ферми зависит от температурного потенциала ϕ_T . Особенно существенна на практике температурная зависимость пикового тока I_1 . Опыт показывает, что эта зависимость может иметь разные знаки при различных концентрациях примесей. Поэтому для ограниченного температурного диапазона можно добиться почти полного постоянства пикового тока. Рабочий диапазон температур, в котором сохраняется отрицательное сопротивление, оказывается весьма широким в связи с использованием вырожденных полупроводников. Для германия максимальная рабочая температура составляет около $+200^\circ\text{C}$, а для арсенида галлия — около 400°C .

Динамические свойства диода ограничены условием сохранения отрицательного сопротивления по отношению к внешней цепи. Рассматривая ту часть эквивалентной схемы диода, которая на рис. 3-10 обведена пунктиром (r_0 — суммарное омическое сопротивление слоев), преобразуя эту часть схемы к виду $R + jX$ и полагая $R = 0$, получаем максимальную рабочую частоту

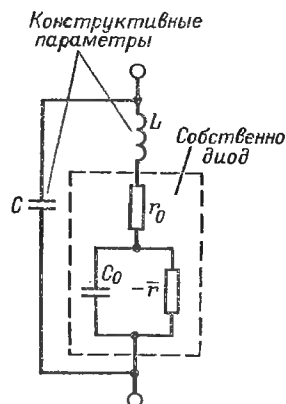


Рис. 3-10. Эквивалентная схема туннельного диода на участке с отрицательным сопротивлением.

¹ Кремниевые туннельные диоды со сколько-нибудь удовлетворительными характеристиками пока получить не удается.

и минимальную постоянную времени в виде

$$f_{\text{макс}} = \frac{1}{2\pi r C_0} \sqrt{\frac{\bar{r}}{r_0} - 1};$$

$$\tau_{\text{мин}} = \frac{\bar{r} C_0}{\sqrt{\bar{r}/r_0 - 1}},$$

где $\bar{r}$ — модуль отрицательного сопротивления; C_0 — барьерная емкость, определяемая выражением (2-826). Поскольку p - n переход в туннельном диоде очень узкий ($l_0 = 0,01 \div 0,015$ мкм), емкость C_0 даже при малой площади (10^{-5} см²) значительна. Если принять типичные значения $r_0 = 1$ Ом; $\bar{r} = 10 \div 100$ Ом и $C_0 = 10 \div 50$ пФ, то $\tau_{\text{мин}} = 0,2 \div 1,0$ нс; $f_{\text{макс}} = 20 - 1000$ МГц, т. е. время переходных процессов может составлять доли наносекунды, а усиливаемые и генерируемые частоты — сотни и тысячи мегагерц.

Следует заметить, что значение $\bar{r}$ не остается постоянным на отрицательном участке: оно минимально при $U \approx 1,8 U_1$ и увеличивается по обе стороны от этой точки. Поэтому частотные свойства диода существенно зависят от режима.

При работе туннельных диодов в импульсных схемах, в которых токи и напряжения меняются в широких пределах, более удобным показателем быстродействия (длительности фронтов) является, как показано в работе [54], отношение I_1/C_0 . Чем больше это отношение, тем выше скорость перезаряда емкости и тем меньше длительность фронтов. Для германиевых диодов указанное отношение обычно составляет десятые доли миллиампера на пикофараду; для арсенидо-галлиевых диодов оно примерно на один порядок выше (1 мА/пФ соответствует скорости изменения напряжения 1 В/нс).

При работе туннельного диода существенную роль играют паразитные индуктивности вводов и тем более подводящих проводов (рис. 3-10). Даже весьма малая индуктивность ($L \approx 0,01$ мкГ) при суммарной емкости схемы 20 пФ соответствует частоте 500 МГц, лежащей в диапазоне рабочих частот диода. Наличие паразитного колебательного контура с отрицательной составляющей сопротивления нередко приводит к самовозбуждению схем на высокой частоте и требует принятия мер по его предотвращению. Поэтому иногда оказывается целесообразной коаксиальная конструкция диода, характерная для многих элементов СВЧ.

Малая величина отрицательного сопротивления сильно осложняет снятие характеристики диода экспериментальным путем, так как требуется источник напряжения с внутренним сопротивлением не более нескольких ом. В противном случае при измерении будут происходить скачки тока и отрицательный участок останется неисследованным. Не меньшие сложности при измерении обусловлены паразитными индуктивностями и самовозбуждением схемы [16].

Туннельный диод представляет собой универсальный прибор, способный выполнять все функции, свойственные активным элементам электронных схем. Поэтому вопросы применения туннельных диодов составили в свое время отдельную область прикладной элект-

троники. В результате выяснилось, что диодные схемы настолько сильно отличаются от триодных, что требуют пересмотра многих привычных методов построения и расчета электронных схем [55]. В связи с прогрессом транзисторной техники и особенно микроэлектроники такой пересмотр структурных основ электронных схем оказался практически нецелесообразным. Соответственно туннельные диоды в настоящее время не являются альтернативой транзисторов, а занимают свое, сравнительно узкое место в области генераторов и усилителей СВЧ диапазона.

Интересным вариантом туннельного диода является так называемый *обращенный диод*, характеристика которого показана на рис. 3-11. Как видим, особенность обращенного диода состоит в том, что отсутствует (или очень мал) максимум на прямой ветви. В этом случае логично повернуть характеристику на 180° (показано пунктиром) и считать прямую ветвь обратной, а обратную — прямой. Таким образом, обращенный диод имеет значительно меньшее прямое напряжение, чем обычные полупроводниковые диоды (у которых оно составляет сотни милливольт и более), что очень ценно для многих применений [16]. Однако его обратное напряжение тоже весьма мало (0,3—0,5 В) и с этим нужно считаться при расчете схем. Технология туннельных и обращенных диодов почти одинакова; различие состоит главным образом в подборе исходных материалов (у обращенных диодов концентрация примесей меньше, чем у туннельных).

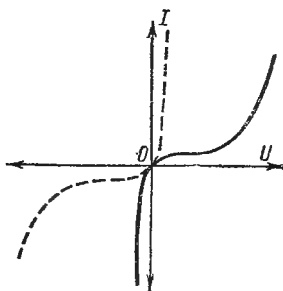


Рис. 3-11. Статическая характеристика обращенного диода.

3-4. ДИОДЫ ШОТТКИ

В основе диодов Шоттки лежит контакт между металлом и полупроводником. Такой контакт, как отмечалось в § 2-4, при определенных условиях может обладать выпрямительными свойствами. Для этого необходимо, чтобы приповерхностный слой полупроводника в равновесном состоянии был обеднен основными носителями и чтобы сопротивление обедненного слоя было много больше сопротивления остальной части полупроводниковой пластины. Несмотря на то что теория контакта между металлом и полупроводником существует около 50 лет ¹, реализовать данный тип диодов удалось лишь в начале 70-х годов. На пути реализации такого, казалось бы, простого прибора стояли следующие трудности: во-первых, нужно было осуществить органический контакт металл — полупроводник

¹ Диоды названы по имени одного из ранних исследователей таких контактов — немецкого физика Шоттки.

без каких бы то ни было промежуточных слоев (прижимный контакт не обеспечивал этого условия); во-вторых, и это главное, нужно было добиться малого сопротивления полупроводниковой пластины, не жертвуя при этом ее удельным сопротивлением (в противном случае резко уменьшилось бы пробивное напряжение, см. § 2-7). Указанные трудности были преодолены путем использования современных технологических методов — вакуумного напыления металлических контактов и эпитаксиального наращивания тонкого высокоомного слоя полупроводника на низкоомной подложке (см. § 4-13).

Основным преимуществом диодов Шоттки по сравнению с диодами с p - n переходом является тот факт, что у них отсутствует инжекция неосновных носителей при прямом смещении, а значит, и явления накопления и рассасывания этих носителей. Соответственно инерционность диодов Шоттки обусловлена только барьерной емкостью контакта и может быть сделана весьма малой путем уменьшения размеров структуры. Типичный диапазон рабочих частот составляет 3—15 ГГц, а времена переключения доходят до 0,1 нс и менее.

Еще одним преимуществом диодов Шоттки является то, что вольт-амперная характеристика (2-33) сохраняется для них в гораздо более широком диапазоне токов, чем для обычного p - n перехода, поскольку отсутствует модуляция сопротивления базы неосновными носителями. Например, логарифмическая зависимость $U(I)$ может выполняться в диапазоне токов от 10^{-12} до 10^{-4} А, т. е. на протяжении восьми декад, что недостижимо в случае p - n переходов. Следовательно, диоды Шоттки могут, по-видимому, использоваться и в качестве логарифмирующих элементов.

Обратный ток диодов Шоттки выражается формулой

$$I_0 = \frac{1}{4} q n_s v_T S,$$

где n_s — граничная концентрация носителей в равновесном состоянии; v_T — средняя тепловая скорость носителей (1-27). Обратные токи могут составлять всего несколько пикоампер. Обратные напряжения лежат в пределах от 10 до 500 В.

Значения прямых токов зависят, конечно, от площади структуры и качества теплоотвода. В настоящее время разработаны диоды Шоттки на прямые токи порядка 50 А. При этом прямое падение напряжения составляет всего 0,5 В, т. е. примерно вдвое меньше, чем у обычных кремниевых диодов.

Основным полупроводником, используемым в диодах Шоттки, является кремний. В качестве металлов используются молибден, никром, золото, платина (точнее, сплав платины с кремнием — силицид платины), а также алюминий — основной металл для омических контактов в интегральных схемах. Концентрация примеси в высокоомном эпитаксиальном слое, как правило, не превышает 10^{17} см⁻³, а толщина этого слоя обычно составляет несколько микрон.

Разумеется, при подборе металла в диодах Шоттки следует учитывать и контактную разность потенциалов φ_{MS} (см. табл. 2-1), и равновесный поверхностный потенциал φ_{so} , обусловленный поверхностными уровнями (см. § 2-4), и такие технологические факторы, как адгезия (степень сцепления) металла с полупроводником, возможность присоединения внешних проволочных выводов и т. п.

Глава четвертая ТРАНЗИСТОРЫ

4-1. ВВЕДЕНИЕ

Плоскостной транзистор, разработанный в 1949—1950 гг. американским физиком В. Шокли [2, 24], — наиболее распространенный тип транзисторов. Он может с успехом выполнять как усилительные, так и ключевые функции, т. е. представляет собой универсальный элемент электронных схем.

Транзистор представляет собой двухпереходный прибор (рис. 4-1). Переходы образуются на границах тех трех слоев, из которых состоит транзистор. Контакты с внешними электродами — омические. В зависимости от типа проводимости крайних слоев различают транзисторы $p-n-p$ и $n-p-n$ со взаимно противоположными рабочими полярностями, что не имеет аналогии в ламповой технике¹. Чтобы не дублировать всех рассуждений и выводов, в дальнейшем будем рассматривать только транзисторы $p-n-p$. Условные обозначения обоих типов транзисторов, рабочие полярности напряжений и направления токов показаны на рис. 4-2.

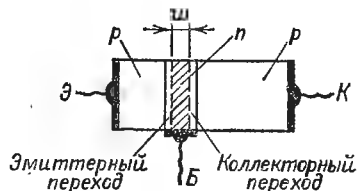


Рис. 4-1. Упрощенная структура транзистора.

Переход, работающий в прямом направлении, называется эмиттерным, а соответствующий крайний слой — *эмиттером*. Такое название, как и у диодов, отражает факт инжекции неосновных носителей через переход. Средний слой называется *базой*. Второй переход, нормально смещенный в обратном направлении, называется коллекторным, а соответствующий крайний слой — *коллектором*.

¹ В транзисторах типа $n-p-n$ рабочими носителями являются электроны и полярности получаются такие же, как у электронных ламп. В транзисторах типа $p-n-p$ рабочими носителями являются дырки и полярности соответствуют воображаемой позитронной лампе. В микроэлектронике (в кремниевых интегральных микросхемах) главную роль играют $n-p-n$ транзисторы. Тот факт, что ниже в основу анализа положены транзисторы $p-n-p$, объясняется чисто историческими причинами: 1-е издание данной книги вышло тогда, когда микроэлектроника только зарождалась, а в электронной технике господствовали германиевые $p-n-p$ транзисторы.

Это название отражает функцию «собираания» инжектированных носителей, прошедших через слой базы. Для того чтобы такое «собираание» было возможно, база должна иметь достаточно малую толщину w . В противном случае инжектированные носители успеют рекомбинировать в процессе перемещения через базу, что мы видели на примере диодов с толстой базой (см. рис. 2-19).

Необходимо подчеркнуть, что транзистор представляет собой, вообще говоря, обратимый прибор, т. е. эмиттер и коллектор можно поменять местами, сохранив в той или иной мере работоспособность прибора. Такой вывод вытекает из однотипности крайних слоев. Однако, в связи с несим-

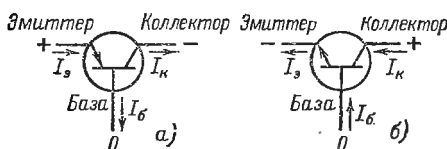


Рис. 4-2. Условные обозначения транзисторов.

а — транзистор $p-n-p$; б — транзистор $n-p-n$.

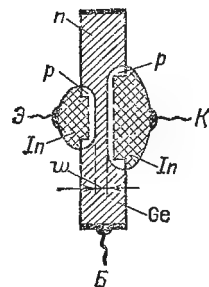


Рис. 4-3. Реальная структура сплавного бездрейфового (диффузионного) транзистора.

метричностью реальной структуры (рис. 4-3, 4-36), а также различием материалов эмиттера и коллектора в большинстве типов транзисторов *нормальное* и *инверсное* включения неравноценны, в чем мы убедимся позднее.

Транзистор иногда работает и в таком своеобразном режиме, когда оба перехода смещены в прямом направлении. При этом, по существу, имеют место двусторонняя инжекция и двустороннее «собираание» неосновных носителей. Если функция инжекции превалирует на обоих переходах, транзистор превращается в двойной диод. Однако чаще на одном из переходов (обычно коллекторном) превалирует функция «собираания», и тогда ток через него протекает в направлении, не соответствующем полярности смещения. Такой режим называется *режимом насыщения* по причинам, которые рассматриваются в гл. 15.

Из всего сказанного следует, что транзистор является системой двух взаимодействующих $p-n$ переходов и что непременным условием такого взаимодействия является достаточно малая толщина базы ($w \ll L$, где L — диффузионная длина неосновных носителей). Основные свойства транзистора определяются процессами в базе, и им в дальнейшем будет уделено главное внимание. Характер движения инжектированных носителей в базе в общем случае заключается в сочетании диффузии и дрейфа. Электрическое поле, в котором происходит дрейф, может быть результатом высокого уровня инжекции (см. § 2-8), а также результатом неоднородности

слоя (см. рис. 1-17, а). Последний случай имеет особенно большое значение, так как собственное поле неоднородного полупроводника обуславливает дрейфовый механизм движения носителей независимо от уровня инжекции. Транзисторы без собственного поля базы называются *диффузионными* или *бездрейфовыми*, а с собственным полем — *дрейфовыми*. Оба названия отражают главный механизм перемещения носителей, хотя, как правило, диффузия и дрейф сочетаются. Более просты для анализа бездрейфовые транзисторы, которым ниже уделено главное внимание. Особенности дрейфовых транзисторов будут рассмотрены в отдельном параграфе.

Выше подразумевалось, что оба напряжения (U_B и U_K) отсчитываются от базы, принятой за *о с н о в н о й* электрод, общий для входной и выходной цепей транзистора. Такое включение транзис-

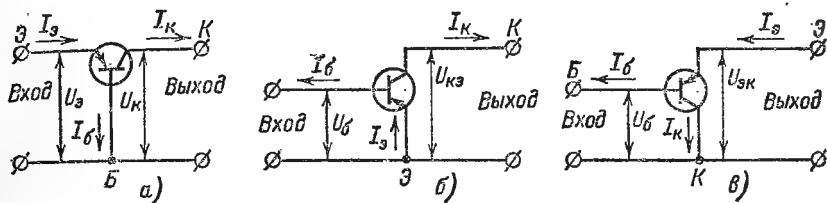


Рис. 4-4. Схемы включения транзистора.

а — с общей базой; б — с общим эмиттером; в — с общим коллектором.

тора (рис. 4-4, а), позволяющее строго и наглядно изучить его физические свойства и параметры, называют включением «с общей базой». Это включение будет в дальнейшем обозначаться буквами ОБ. Однако схема ОБ является не только не единственно возможной, но даже не наиболее распространенной на практике. Это объясняется рядом обстоятельств (например, отсутствием усиления тока), которые будут ясны из последующих глав. Основное применение в схемах находит другое включение транзистора, которое по вполне понятным причинам называют включением «с общим эмиттером» (рис. 4-4, б). Мы будем обозначать его буквами ОЭ. Преимущества схемы ОЭ также выяснятся в дальнейшем; однако можно сразу отметить, что она дает усиление по току, поскольку ток базы, являющийся для нее входным, гораздо меньше токов эмиттера и коллектора. Третий вариант включения — схема «с общим коллектором» (ОК) — показан на рис. 4-4, в.

Несмотря на практические недостатки, схема ОБ является основой при рассмотрении физических процессов в транзисторе. Поэтому именно она будет предметом анализа в ближайших параграфах.

4-2. ОСНОВНЫЕ ПРОЦЕССЫ В ТРАНЗИСТОРЕ

На рис. 4-3 показан разрез бездрейфового транзистора со сплавными переходами (см. § 4-13), имеющего дисковую структуру. По сравнению с другими типами транзисторов сплавной транзистор

наиболее прост и удобен для анализа. База этого транзистора однородна, поэтому механизм движения носителей — диффузионный. Удельные сопротивления слоев эмиттера и коллектора практически одинаковы, так что фактором, обуславливающим асимметрию транзистора, является только различие площадей S_e и S_k . Асимметрия транзистора преследует ту цель, чтобы дырки, инжектируемые эмиттером и диффундирующие под некоторым углом к оси транзистора, по возможности полнее собирались коллектором.

База сплавного транзистора отличается от базы идеализированной структуры (см. рис. 4-1) наличием трех участков, которые называют активной, промежуточной (или коллекторной) и пассивной областями базы. Активной областью базы является цилиндрический объем с высотой w и площадью, равной поверхности эмиттера S_e . Промежуточной областью базы является кольцевой объем с площадью основания $S_k - S_e$ и высотой, равной расстоянию от коллектора до противоположной поверхности базовой пластинки. Наконец, пассивной областью базы является ее объем, расположенный вне коллектора. Для первоначального ознакомления с транзистором можно пренебречь пассивной и промежуточной областями и считать транзистор симметричным, имеющим одинаковую площадь $S = S_e$ во всех сечениях (см. рис. 4-1). Обычно соблюдается соотношение $w \ll \sqrt{S}$, т. е. размеры транзистора в направлениях, перпендикулярных главной оси, много больше толщины базы. В таком транзисторе краевые эффекты не очень существенны, и его можно считать одномерным, т. е. можно предположить движение носителей только вдоль главной оси, без отклонения в стороны. Такая одномерная модель будет всегда подразумеваться, если не сделано специальных оговорок¹.

• Инжекция и сборка неосновных носителей. На рис. 4-5, а показана энергетическая диаграмма транзистора в равновесном состоянии. Диаграмма показывает, что эмиттер и коллектор представляют собой низкоомные слои (уровень Ферми лежит вблизи уровней акцепторов), а база — сравнительно высокоомный слой (уровень Ферми расположен вблизи середины запрещенной зоны). Легко видеть, что электроны базы и дырки эмиттера и коллектора находятся в «потенциальных ямах», из которых они могут перейти в смежный слой только благодаря достаточно большой тепловой энергии. Наоборот, дырки базы и электроны эмиттера и коллектора находятся на «потенциальных гребнях», с которых они могут свободно переходить в смежный слой. В равновесном состоянии на обоих переходах имеется динамическое равновесие между потоками дырок (а также между потоками электронов), протекающих в ту и другую сторону.

¹ Одномерная модель не может отразить влияния поверхностной рекомбинации, но это влияние, как известно, учитывается при использовании эффективного времени жизни (см. § 1-10).

Пусть на транзитерном переходе задано нормальное для него положительное смещение, а коллекторный переход по-прежнему замкнут (рис. 4-5, б). Тогда потенциальный барьер эмиттера понизится и начнется инжекция дырок в базу и электронов в эмиттер.

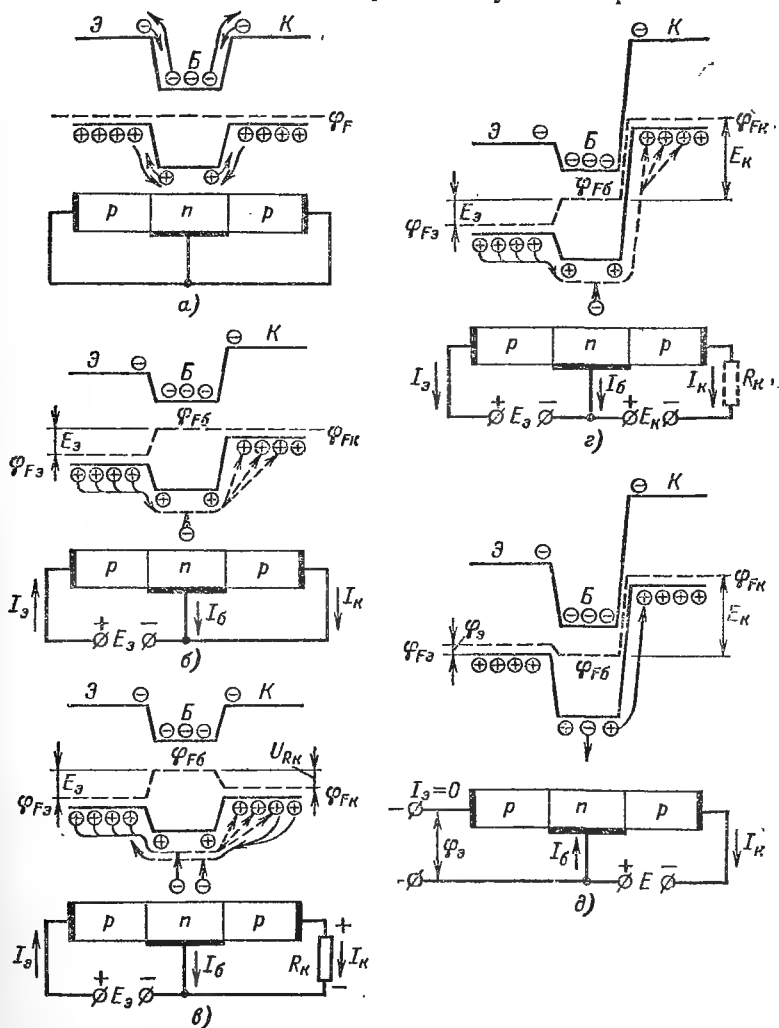


Рис. 4-5. Зонные диаграммы для транзистора при различных режимах его работы.

При большой разнице в удельных сопротивлениях слоев Э и Б электронная составляющая тока, как известно, не играет большой роли и ею пока можно пренебречь. Инжектированные дырки, пройдя базу, доходят до коллекторного перехода и свободно проходят в коллектор. Значит, в выходной цепи будет протекать ток, близкий

к току эмиттера, поскольку рекомбинация в тонкой базе невелика. Небольшая разность между эмиттерным и коллекторным токами составляет ток базы, причем этот ток обусловлен электронами, которые призваны пополнять убыль электронов в базе при рекомбинации их с дырками. Поскольку напряжение U_k равно нулю, полезная мощность не выделяется и усиление отсутствует. Если в выходную цепь включить сопротивление для выделения мощности (рис. 4-5, в), то падение напряжения на этом сопротивлении создаст положительное смещение коллектора. Тогда наряду с собиранием дырок, дошедших до эмиттера, будет происходить инжекция дырок самим коллектором. В результате коллекторный ток станет заметно меньше тока эмиттера, мощность в нагрузке будет очень невелика и у с и л е н и я мощности не будет.

В нормальном усилительном режиме на коллекторный переход задается достаточно большое отрицательное смещение, которое приводит к существенному повышению потенциального барьера у коллектора (рис. 4-5, г). Теперь можно включать в выходную цепь значительные сопротивления без опасения вызвать инжекцию через коллекторный переход¹. При этом можно получить значительную выходную мощность, а главное — усиление мощности, так как токи I_e и I_k почти одинаковы, а сопротивление нагрузки превышает сопротивление эмиттерного перехода.

Пусть теперь коллекторный переход смещен в обратном направлении, а эмиттер «оборван» (рис. 4-5, д). Высокий потенциальный барьер коллектора практически исключает уход дырок из коллектора в базу. Следовательно, ток через коллекторный переход обусловлен неуравновешенным потоком дырок из базы в коллектор. Токи I_k и I_b в этом случае невелики и, конечно, равны друг другу, поскольку $I_e = 0$. Экстракция дырок из базы через коллекторный переход создает отрицательный градиент их концентрации вдоль базы в сторону коллектора. В результате уменьшается тот поток дырок из базы в эмиттер, который был в равновесном состоянии (рис. 4-5, а), и встречный поток дырок из эмиттера в базу оказывается неуравновешенным. Это вызывает накопление положительного заряда в базе, а в эмиттере образуется такой же отрицательный заряд. Соответственно несколько возрастает разность потенциалов на эмиттерном переходе, так что поток дырок из эмиттера уменьшается и в конце концов снова становится равным потоку дырок из базы. Тогда ток I_e делается равным нулю, как и должно быть при оборванном эмиттере.

Аналогично можно рассмотреть и другие возможные режимы транзистора, в том числе инверсное включение (см. § 4-1).

Во всех случаях у транзистора $p-n-p$ главными рабочими носителями, образующими токи через переходы, являются дырки, а ток базы всегда обусловлен электронами; последние ком-

¹ Для этого сопротивление должно удовлетворять условию

$$-E_k + I_k R_k < 0.$$

пенсируют избыточный заряд дырок в базе и обеспечивают ее нейтральность как во время переходных процессов (когда дырки поступают или уходят из базы), так и в стационарном режиме (когда убыль дырок обусловлена только рекомбинацией).

Распределение носителей в базе. Дырки, инжектированные эмиттером, достигают коллектора не сразу, а с некоторой задержкой, обусловленной их перемещением вдоль базы. Кроме того, в связи с хаотичностью движения дырок коллекторный ток нарастает не скачком, а плавно. Соответственно ток базы в первый момент равен току эмиттера, а затем постепенно уменьшается до стационарного значения. Примерная картина переходного процесса показана на рис. 4-6, где t_s — время задержки, а t_D — среднее время диффузии.

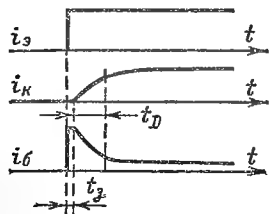


Рис. 4-6. Переходные процессы при подаче ступеньки эмиттерного тока.

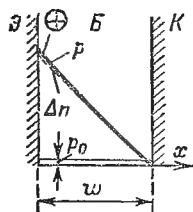


Рис. 4-7. Распределение дырок и избыточных электронов в базе.

В установившемся режиме дырки в базе транзистора распределены почти так же, как в диоде с тонкой базой. Это объясняется сходством граничных условий. В самом деле, для диода было принято условие (2-286): $\Delta p(w) = 0$; для транзистора при отрицательном смещении коллектора согласно (2-14а) получаем: $\Delta p(w) \approx -p_0$. Поскольку обычно $p_0 \ll \Delta p(0)$, можно считать эти граничные условия практически одинаковыми. Тогда, используя выражение (2-31б) (так как $w \ll L$), приходим к выводу, что стационарное распределение дырок в базе почти линейно, как показано на рис. 4-7. На самом деле градиент концентрации около коллекторного перехода несколько меньше, чем около эмиттерного, поскольку ток коллектора (из-за рекомбинации) немного меньше эмиттерного тока. Эту разницу в градиентах следует иметь в виду, но ее трудно отразить на графике. Необходимо подчеркнуть, что линейное распределение свойственно только стационарному режиму. Во время коротких переходных процессов (длительностью порядка t_D и менее) распределение может существенно отличаться от линейного (см. рис. 15-16 и 15-18).

Линейному распределению дырок должно соответствовать почти линейное распределение компенсирующих (избыточных) электронов в базе (рис. 4-7). Причины небольшого различия в распределении дырок и избыточных электронов те же, что и в случае диода (см. § 2-8).

Заряды избыточных носителей. пропорциональны площадям под кривыми их распределения. Поскольку база в целом нейтральна, можно считать эти площади одинаковыми. Для оценки заряда удобнее пользоваться распределением дырок. Очевидно, что заряд дырок пропорционален толщине базы и току транзистора, определяющему наклон линии $p(x)$. Вопрос о том, какому из двух токов (I_s или I_k) пропорционален заряд, не очень существен, так как эти токи в стационарном режиме почти одинаковы. В большинстве случаев удобнее считать заряд пропорциональным току коллектора (см. § 15-3), так как в схемах этот ток обычно не претерпевает скачкообразных изменений.

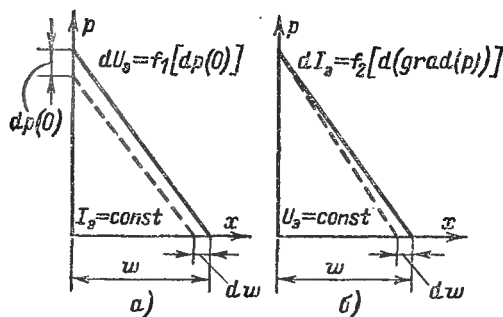


Рис. 4-8. Влияние модуляции толщины базы на входные величины.

$a - I_s = \text{const}$; $b - U_s = \text{const}$.

Эквивалентную емкость, обусловленную изменениями заряда в базе, называют, как и в диоде, диффузионной емкостью (см. § 2-9). Модуляция толщины базы. Как известно, ширина $p-n$ перехода зависит от напряжения на нем. Поскольку эмиттерный переход смещен в прямом направлении, его ширина мала и изменения этой ширины при изменениях U_s не имеют существенного значения. Коллекторный же переход, смещенный в обратном направлении, имеет сравнительно большую ширину, и изменения ее при изменениях напряжения U_k важны для работы транзистора. А именно, поскольку коллекторный переход сосредоточен в базе (как более высокоомном слое), приращения его ширины вызывают практически равные им приращения толщины базы. В результате получается зависимость $w(U_k)$, которую называют *модуляцией толщины базы* или *эффектом Эрли* [56]. Посмотрим, каких следствий можно ожидать от этого эффекта.

Во-первых, изменение толщины базы влияет на ту долю инжектированных дырок, которая доходит до коллектора, избежав рекомбинации. Чем меньше толщина базы, тем эта доля больше. Значит, при неизменном токе эмиттера модуляция толщины базы приводит к изменениям тока коллектора. Соответственно коэффициент передачи эмиттерного тока оказывается функцией коллекторного напряжения, а коллекторный переход имеет конечное дифференциальное сопротивление.

Во-вторых, модуляция толщины базы сопровождается изменением заряда дырок в базе; иначе говоря, имеет место зависимость заряда от коллекторного напряжения, т. е. коллекторный пере-

ход обладает некоторой диффузионной емкостью дополнительно к обычной барьерной.

В-третьих, модуляция толщины базы меняет время диффузии дырок через базу; тем самым коллекторное напряжение влияет на частотные свойства транзистора.

В-четвертых, поскольку тепловой ток эмиттерного перехода $I_{\text{эо}}$ при тонкой базе обратно пропорционален ее толщине [см. (2-366)], напряжение $U_{\text{к}}$, модулируя толщину базы, модулирует также ток $I_{\text{эо}}$, а вместе с ним согласно (2-33) всю вольт-амперную характеристику эмиттерного перехода. Следовательно, если одна из входных величин ($I_{\text{э}}$ или $U_{\text{э}}$) задана, то вторая оказывается функцией коллекторного напряжения (рис. 4-8)¹. Такое влияние разумно назвать *внутренней обратной связью по напряжению*.

4.3. СТАТИЧЕСКИЕ ХАРАКТЕРИСТИКИ ТРАНЗИСТОРА

Выше была рассмотрена идеализированная модель транзистора. Идеализация заключалась не только в том, что модель считалась одномерной, но и в том, что не учитывались объемные сопротивления слоев. Сопротивления слоев эмиттера и коллектора существенны только в некоторых ключевых режимах (см. гл. 15). Сопротивление же базы существенно почти во всех случаях, но, чтобы не усложнять предварительный анализ, будет учтено позднее.

Формулы Молла — Эберса. Приступая к выводу основных характеристик, примем еще одно упрощение, а именно пренебрежем эффектом модуляции толщины базы вместе с его следствиями. Тогда для транзистора можно принять такую эквивалентную схему [57], которая показана на рис. 4-9. Здесь каждый из переходов изображен в виде диода, а взаимодействие их отражено генераторами токов.

Так, если эмиттерный переход открыт и через него протекает ток I_1 , то в цепи коллектора, как известно, будет протекать несколько меньший ток, поскольку часть инжектированных носителей рекомбинирует. Этот меньший ток обеспечивается на схеме генератором $\alpha_N I_1$, где $\alpha_N < 1$ — коэффициент передачи эмиттерного тока. Индекс N означает нормальное включение транзистора. Если триод работает в инверсном включении (положительное смещение на коллекторе и отрицательное на эмиттере), то прямому

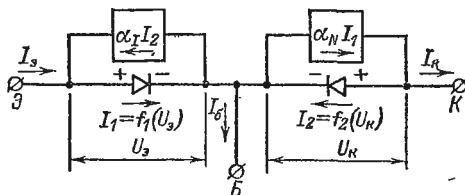


Рис. 4-9. Эквивалентная схема идеализированного транзистора.

¹ В общем случае, когда источник смещения в цепи эмиттера имеет конечное сопротивление, изменения $U_{\text{к}}$ влияют и на ток $I_{\text{э}}$, и на напряжение $U_{\text{э}}$.

коллекторному току I_2 соответствует эмиттерный ток $\alpha_I I_2$, вытекающий из эмиттера. Коэффициент α_I есть коэффициент передачи коллекторного тока, а индекс I означает инверсное включение¹. Таким образом, в общем случае токи эмиттера и коллектора складываются из двух компонентов: инжектируемого (I_1 или I_2) и собираемого ($\alpha_I I_2$ или $\alpha_N I_1$):

$$I_3 = I_1 - \alpha_I I_2; \quad (4-1a)$$

$$I_K = \alpha_N I_1 - I_2. \quad (4-1b)$$

Связь инжектируемых компонентов с напряжениями на переходах такая же, как и в отдельном диоде, т. е. в простейшем случае выражается формулой (2-33)²:

$$I_1 = I'_{30} \left(e^{\frac{U_3}{\Phi_T}} - 1 \right); \quad (4-2a)$$

$$I_2 = I'_{K0} \left(e^{\frac{U_K}{\Phi_T}} - 1 \right). \quad (4-2b)$$

Здесь I'_{30} и I'_{K0} — тепловые токи эмиттерного и коллекторного диодов, измеряемые соответственно при $U_K = 0$ и $U_3 = 0$. Тепловые токи I'_{30} и I'_{K0} выразим через такие величины, которые обычно задаются в технической документации на транзистор, а именно через токи I_{30} и I_{K0} , измеряемые соответственно при $I_K = 0$ и $I_3 = 0$.

Оборвем цепь эмиттера и подадим на оставшийся «коллекторный диод» достаточно большое запирающее напряжение $|U_K| \gg \Phi_T$. Коллекторный ток, который при этом будет протекать, обозначим через I_{K0} и назовем *тепловым током коллектора* в соответствии с терминологией для диодов. Происхождение этого тока было рассмотрено в связи с рис. 4-5, д. Теперь легко выразить ток I'_{K0} через ток I_{K0} . Из формулы (4-1a) при $I_3 = 0$ получаем: $I_1 = \alpha_I I_2$; из формулы (4-2b) при $|U_K| \gg \Phi_T$ получаем: $I_2 = -I'_{K0}$. Подставляя эти значения в (4-1b) и полагая $I_K = I_{K0}$, получаем:

$$I'_{K0} = \frac{I_{K0}}{1 - \alpha_N \alpha_I}. \quad (4-3a)$$

Обозначив ток эмиттера при большом отрицательном смещении ($|U_3| \gg \Phi_T$) и оборванном коллекторе через I_{30} (*тепловой ток эмиттера*), аналогичным путем получим³:

$$I'_{30} = \frac{I_{30}}{1 - \alpha_N \alpha_I}. \quad (4-3b)$$

¹ Различая коэффициенты передачи (а в дальнейшем и тепловые токи) для нормального и инверсного включений, мы косвенно учитываем неоднородность транзистора — его асимметрию, хотя в каждом из включений считаем движение носителей по-прежнему одномерным.

² Более общие формулы (2-73) и (2-79б) действительны только при прямых смещениях.

³ Поскольку ток I_{30} вытекает из эмиттера, в выражении (4-1a) следует считать $I_3 = -I_{30}$.

Подставив токи I_1 и I_2 из (4-2) в соотношения (4-1), найдем зависимости I_9 (U_9 ; U_K) и I_K (U_9 ; U_K), т. е. статические вольт-амперные характеристики транзистора:

$$I_9 = I'_{90} \left(e^{\frac{U_9}{\varphi_T}} - 1 \right) - \alpha_I I'_{K0} \left(e^{\frac{U_K}{\varphi_T}} - 1 \right); \quad (4-4a)$$

$$I_K = \alpha_N I'_{90} \left(e^{\frac{U_9}{\varphi_T}} - 1 \right) - I'_{K0} \left(e^{\frac{U_K}{\varphi_T}} - 1 \right). \quad (4-4b)$$

Запишем еще ток базы, равный разности токов I_9 и I_K :

$$I_6 = (1 - \alpha_N) I'_{90} \left(e^{\frac{U_9}{\varphi_T}} - 1 \right) + (1 - \alpha_I) I'_{K0} \left(e^{\frac{U_K}{\varphi_T}} - 1 \right). \quad (4-4в)$$

Формулы Молла — Эберса (4-4), несмотря на их приближенность, очень полезны для анализа статических режимов, так как хорошо отражают основные особенности транзисторов при любых сочетаниях напряжений на переходах¹.

Можно показать, что в транзисторах выполняется соотношение

$$\alpha_N I_{90} = \alpha_I I_{K0}, \quad (4-5)$$

которое позволяет упрощать формулы (4-4) и выводы из них. В частности, поскольку значения α_N и α_I различаются не очень сильно, в первом приближении можно полагать $I_{90} \approx I_{K0}$, хотя в принципе $I_{90} < I_{K0}$.

Идеализированные статические характеристики. В гл. 2 было показано, что задать прямое напряжение на p - n переходе трудно. Поэтому в большинстве случаев целесообразно считать заданной величиной эмиттерный ток, а не эмиттерное напряжение. Выражая двучлен $e^{U_9/\varphi_T} - 1$ из формулы (4-4a) и подставляя его в (4-4b), получаем:

$$I_K = \alpha_N I_9 - I_{K0} (e^{U_K/\varphi_T} - 1). \quad (4-6)$$

Это выражение представляет собой семейство коллекторных характеристик I_K (U_K) с параметром I_9 . Такое семейство показано на рис. 4-10, а. Семейство эмиттерных характеристик U_9 (I_9) с параметром U_K получается из выражения (4-4a), если разрешить его относительно U_9 . Используя соотношение (4-5), получаем:

$$U_9 = \varphi_T \ln \left[\frac{I_9}{I'_{90}} + 1 + \alpha_N (e^{U_K/\varphi_T} - 1) \right]. \quad (4-7)$$

Эмиттерное семейство характеристик показано на рис. 4-10, б.

¹ Применительно к кремниевым транзисторам нужно иметь в виду следующие особенности: 1) формулы (4-4) недействительны для режима отсечки, так как тепловые токи не являются главными составляющими обратных токов кремниевых переходов (см. § 2-6); 2) соответственно тепловые токи нельзя измерить тем методом, который описан выше; их рассчитывают по прямой характеристике, измеряя токи и напряжения с высокой точностью; 3) при токе базы от 1 мкА и выше в формуле (4-4в) следует заменить φ_T на $1,2\varphi_T$.

Из рис. 4-10, а ясно видны два резко различных режима работы транзистора: *активный режим*, соответствующий значениям $U_K < 0$ (первый квадрант), и *режим насыщения*, соответствующий значениям $U_K > 0$ (второй квадрант). Активный режим является основным в усилительной технике и будет подробно изучен в последующих параграфах. Режим насыщения характерен для ключевых импульсных схем и будет рассмотрен в гл. 15.

Для активного режима характерны условия $U_K < 0$ и $|U_K| \gg \varphi_T$, при которых формулы (4-6) и (4-7) переходят в следующие:

$$I_K = \alpha I_B + I_{K0}; \quad (4-8)$$

$$U_B = \varphi_T \ln \frac{I_B}{I'_{B0}}. \quad (4-9)$$

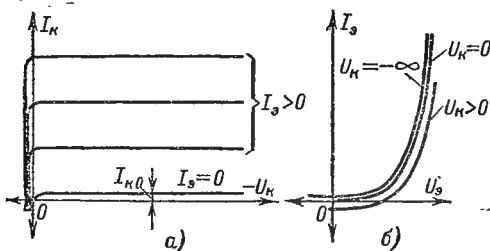


Рис. 4-10. Статические характеристики идеализированного транзистора.

а — коллекторные; б — эмиттерные.

В формуле (4-8), широко используемой на

практике, для простоты опущен индекс N при коэффициенте α , а при выводе формулы (4-9) для простоты положено $1 - \alpha_N = 0$, что вполне оправдано, если $I_B \gg I'_{B0}$.

Характеристики на рис. 4-10, а являются эквидистантными. Эквидистантность характеристик обусловлена принятым при построении постоянством параметра α . Реальные характеристики, как увидим позднее, неэквидистантны, так как α зависит от тока. Кроме того, реальные характеристики имеют конечный наклон, обусловленный не учтенным в формулах (4-4) сопротивлением коллекторного перехода (следствие модуляции толщины базы).

Относительно эмиттерного семейства (рис. 4-10, б) можно сделать следующие замечания. Кривая с параметром $U_K = 0$, естественно, является обычной диодной характеристикой. При значениях $U_K > 0$ кривые сдвигаются вправо и вниз в связи с нарастанием собираемого компонента эмиттерного тока. При малых значениях $U_K < 0$ кривые очень незначительно смещаются влево и вверх. Если же $|U_K| \gg \varphi_T$, то влияние коллекторного напряжения практически отсутствует. На реальных характеристиках, как увидим ниже, влияние отрицательного напряжения U_K тоже невелико, но все же имеет место при любых значениях U_K из-за внутренней обратной связи по напряжению (следствие модуляции толщины базы).

Реальные статические характеристики. В формулах Молла — Эберса не учитывается целый ряд факторов, таких, как эффект Эрли, пробой перехода, зависимость α от тока и др. Поэтому характеристики на рис. 4-10 в значительной степени идеализированы. Реальные коллекторные и эмиттерные характеристики показаны на рис. 4-11. Кривые коллекторного семейства имеют конечный;

хотя и очень небольшой, наклон, который в области, близкой к пробою, резко увеличивается. Расстояние между кривыми немного уменьшается при больших токах из-за уменьшения α . На рис. 4-11, а проведена гипербола допустимой мощности, рассеиваемой в основном на коллекторном переходе. При нагреве транзистора кривые смещаются вверх, в область больших токов, из-за роста тока $I_{к0}$.

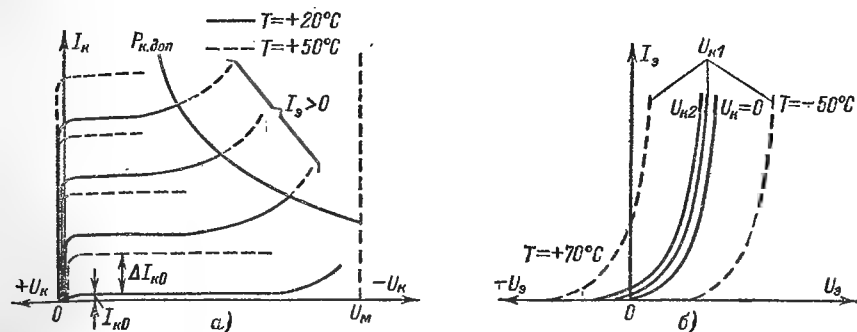


Рис. 4-11. Статические характеристики транзистора при включении по схеме с общей базой.

а — выходные; б — входные.

В активном режиме (1-й квадрант), усредняя нелинейное сопротивление r_k , можно характеризовать коллекторное семейство ОБ достаточно строгим соотношением

$$I_K = \alpha I_E + I_{K0} + \frac{U_K}{r_K}. \quad (4-10)$$

Последний член его обычно несуществен, и в большинстве случаев пользуются упрощенным выражением (4-8). У кремниевых транзисторов (а также у германиевых при не высоких температурах) вполне допустимо пренебречь током I_{K0} , что сильно упрощает расчеты.

Кривые эмиттерного семейства (рис. 4-11, б) образуют довольно плотный «пучок», потому что влияние коллекторного напряжения на эмиттерное (внутренняя обратная связь — следствие эффекта Эрли) очень мало.

При нагреве транзистора кривые смещаются влево в область меньших напряжений. При одном и том же эмиттерном токе эмиттерные напряжения у кремниевых транзисторов на 0,4 В больше, чем у германиевых (см. рис. 2-21) и обычно составляют 0,7—0,8 В. При достаточно большом токе входные вольт-амперные характеристики деформируются — вырождаются из-за падения напряжения в слое базы [см. (2-64), (2-65) и рис. 2-32]. Однако при прочих равных условиях вырождение у транзисторов наступает при значительно больших токах, чем у диодов, так как по сопротивлению базы протекает ток I_6 , в десятки раз меньший тока I_E .

Рассмотрим особенности выходных характеристик транзистора в области пробоя. Если любой из переходов транзистора использовать в качестве диода, то пробой при обратном напряжении будет иметь такой же характер, как в изолированном переходе (см. § 2-7). Например, при оборванном эмиттере коллекторный ток будет равен $MI_{к0}$ (где M — коэффициент ударной ионизации), а напряжением лавинного пробоя будет величина U_M , при которой $M = \infty$. При любом конечном токе эмиттера соответствующий ток I_k тоже увеличивается в M раз, напряжение пробоя остается равным U_M .

Заметим, что чисто лавинный пробой имеет место только при достаточно быстром повышении коллекторного напряжения, когда температура перехода не успевает заметно измениться. В противном случае лавинный пробой может уступить место тепловому, при котором напряжение пробоя будет уменьшаться

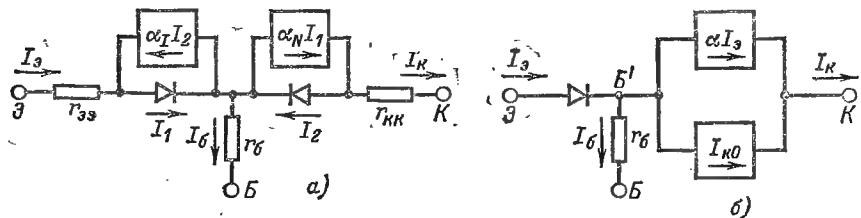


Рис. 4-12. Эквивалентные схемы транзистора для постоянных составляющих.
а — общая схема с учетом сопротивлений слоев; б — схема для нормального активного режима.

с ростом эмиттерного и коллекторного токов. Такой вывод следует из выражения (2-60): чем больше ток I_k , тем выше температура коллекторного перехода, а значит, больше начальное значение теплового тока I_{01} .

Практическая ценность эквивалентной схемы на рис. 4-9 значительно повышается, если дополнить ее сопротивлениями слоев $r_б$, $r_ээ$, $r_кк$ (рис. 4-12, а). Такое дополнение особенно целесообразно тогда, когда один или оба тока $I_э$ и I_k могут считаться заданными. Пусть, например, транзистор работает в активном режиме, т. е. заданы $I_э > 0$ и $U_{кб} < 0$. Ту часть напряжения $U_{кб}$, которая приходится на коллекторный переход, будем считать достаточно большой: $|U_k| \gg \phi_T$. В этом случае согласно (4-26) имеем $I_2 = -I'_{к0}$. Соответственно токи генераторов на рис. 4-12, а будут равны $-\alpha_I I'_{к0}$ и $\alpha_N (I_э - \alpha_I I'_{к0})$. Первым из них для простоты пренебрежем (это вполне допустимо, если $I_э \gg \alpha_I I'_{к0}$), а постоянную составляющую второго ($-\alpha_N \alpha_I I'_{к0}$) объединим с током $I'_{к0}$, также протекающим через коллекторный диод. В результате, учитывая (4-3а), получим эквивалентную схему, показанную на рис. 4-12, б и соответствующую выражению (4-8). Такая схема полезна для расчета режима усилительных каскадов. Сопротивления слоев коллектора и эмиттера здесь опущены, так как в усилительной

технике они несущественны. Однако включение этих сопротивлений в схему не приводит ни к каким затруднениям, поскольку через них протекают заданные токи и, значит, соответствующие напряжения легко рассчитать и добавить к напряжениям на переходах.

4.4. СТАТИЧЕСКИЕ ПАРАМЕТРЫ ТРАНЗИСТОРА

Нелинейные эквивалентные схемы, показанные на рис. 4-12, используются при анализе вопросов, связанных с большим сигналом (см., например, гл. 15). При расчете малых переменных составляющих, характерных для усилительной техники, эти схемы целесообразно линеаризовать.

Возьмем за основу схему на рис. 4-12, б. Генератор постоянного тока $I_{к0}$ исключим, поскольку нас интересуют переменные составляющие, и введем вместо него дифференциальное сопротивление коллекторного перехода r_k . Эмиттерный диод также заменим его дифференциальным сопротивлением r_e . Обратную связь по напряжению (см. конец § 4-2) отразим генератором э. д. с. $\mu_{эк} U_k$, включенным последовательно с сопротивлением r_e . Наконец, для учета частотных зависимостей включим параллельно сопротивлениям r_e и r_k барьерные емкости, а коэффициент α будем считать операторной или комплексной величиной¹. Тогда линейная эквивалентная схема транзистора будет такой, как показано на рис. 4-13. Ее легко дополнить паразитными емкостями, однако необходимость в этом возникает редко.

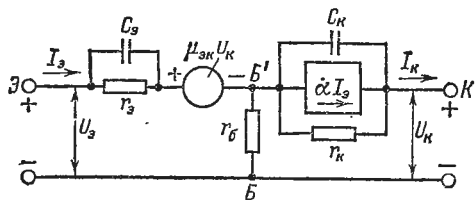


Рис. 4-13. Эквивалентная схема транзистора для переменных составляющих.

Схема на рис. 4-13 хорошо отражает структуру транзистора и содержит физически обоснованные параметры. Точка B' на схеме называется внутренней базовой точкой в отличие от внешнего зажима базы B .

К числу основных параметров, необходимых при построении эквивалентной схемы транзистора (для переменных составляющих), относятся следующие:

1. Дифференциальный коэффициент передачи эмиттерного тока при нормальном включении α_N . В дальнейшем будем писать его без индекса N , так как инверсное включение относится к специаль-

¹ Заметим, что коэффициент α в схеме для переменных составляющих имеет несколько другое значение, чем в схеме для постоянных составляющих (см. ниже).

ным случаям. Коэффициент α определяется следующим образом:

$$\alpha = \left(\frac{dI_K}{dI_9} \right)_{U_K = \text{const}} \quad (4-11a)$$

2. Дифференциальное сопротивление эмиттерного перехода

$$r_9 = \left(\frac{dU_9}{dI_9} \right)_{U_K = \text{const}} \quad (4-11б)$$

3. Дифференциальное сопротивление коллекторного перехода

$$r_K = \left(\frac{dU_K}{dI_K} \right)_{I_9 = \text{const}} \quad (4-11в)$$

4. Коэффициент внутренней обратной связи по напряжению, характеризующий влияние коллекторного напряжения на эмиттерное в связи с модуляцией толщины базы,

$$\mu_{ЭК} = \left(\frac{dU_K}{dU_9} \right)_{I_9 = \text{const}} \quad (4-11г)$$

5. Объемное сопротивление базы r_6 . В отличие от предыдущих параметров сопротивление базы должно определяться не для одномерной модели, а для реальной структуры транзистора, так как ток базы протекает в направлении, перпендикулярном потоку дырок, и, следовательно, необходимо учитывать реальную конфигурацию базы как в активной, так и в пассивной ее части.

Помимо перечисленных дифференциальных параметров важную роль в работе транзистора играет тепловой ток I_{K0} , который определяется следующим образом:

$$I_{K0} = (I_K)_{I_9 = 0; |U_K| \gg \varphi_T}$$

Параметры, характеризующие переходные и частотные свойства транзистора, будут определены в следующем параграфе.

Коэффициент передачи эмиттерного тока. Величина α , стоящая в формуле (4-8), в отличие от величины α в формуле (4-11a) является интегральной, так как связывает не приращения dI_K и dI_9 , а полные токи I_K и I_9 . Снабдив интегральный коэффициент передачи черточкой сверху (чтобы отличить его от дифференциального), получим:

$$\bar{\alpha} = \frac{I_K - I_{K0}}{I_9} \quad (4-12a)$$

Если бы коэффициент α не зависел от тока I_9 , то, как следует из (4-8), дифференциальный коэффициент передачи был бы равен интегральному: $\alpha = \bar{\alpha}$. На самом деле $\bar{\alpha}$ является функцией эмиттерного тока, и поэтому, продифференцировав (4-8) по току I_9 , получим следующее соотношение:

$$\alpha = \bar{\alpha} + I_9 \frac{d\bar{\alpha}}{dI_9} \quad (4-12б)$$

Как видим, α может быть больше или меньше, чем $\bar{\alpha}$, в зависимости от знака производной $d\bar{\alpha}/dI_0$. Соотношение (4-126) позволяет вычислить α , если известна функция $\bar{\alpha}(I_0)$, но практически коэффициенты α и $\bar{\alpha}$ измеряются раздельно.

Наша ближайшая задача состоит в том, чтобы выразить коэффициент передачи тока через физические параметры транзистора. Очевидно, что в математическом плане эта задача будет линейной только в том случае, если пренебречь зависимостью $\bar{\alpha}(I_0)$ и тем самым считать $\alpha = \bar{\alpha}$. Поэтому в дальнейшем мы будем различать интегральный и дифференциальный коэффициенты передачи лишь тогда, когда это принципиально необходимо.

Будем считать, что ток коллектора обусловлен только дырками, доходящими от эмиттера, так как потоки электронов из базы и дырок из коллектора предотвращаются высоким потенциальным барьером в коллекторном переходе (см. рис. 4-5, 2), а поток электронов из коллектора малосуществен, поскольку коллекторный слой — дырочный и концентрация электронов в нем невелика. Что касается тока эмиттера, то он, вообще говоря, обусловлен обоими типами носителей.

Таким образом, запишем коэффициент передачи эмиттерного тока в следующем виде:

$$\alpha = \gamma \kappa, \quad (4-13)$$

где γ — коэффициент инжекции дырок [см. (2-35) и (2-74)]; κ — коэффициент переноса дырок через базу, показывающий, какая доля инжектированных дырок доходит до коллектора¹.

На низких частотах коэффициент инжекции в транзисторах не имеет специфики по сравнению с диодами. Поэтому прежде всего проанализируем коэффициент переноса. Для этого решим уравнение диффузии (1-79а), полагая $\partial p/\partial t = 0$. В этом случае уравнение будет таким же, как (2-26), но мы запишем его для полной концентрации, подставив $\Delta p = p - p_0$; тогда

$$\frac{d^2 p}{dx^2} - \frac{p}{L^2} = -\frac{p_0}{L^2}. \quad (4-14)$$

Общее решение этого уравнения будет таким же, как (2-27), а частным решением будет p_0 , поэтому

$$p(x) = A_1 e^{x/L} + A_2 e^{-x/L} + p_0. \quad (4-15)$$

Граничные условия запишем, исходя из того, что в эмиттерной цепи задан дырочный ток $I_{ep} = \gamma I_0$, а на коллекторном переходе — напряжение U_k . Учитывая (1-73а) и (2-13а), получаем

¹ В физической литературе для коэффициента переноса часто используется обозначение β , которое мы, однако, сохраним для другого параметра (коэффициента передачи тока базы), более важного в транзисторных схемах.

при $x = 0$ и $x = w$ соответственно:

$$-\frac{dp}{dx} \Big|_{x=0} = \frac{I_{\text{эп}}}{qDS}; \quad (4-16a)$$

$$p(w) = p_0 e^{\frac{U_K}{\Psi_T}}, \quad (4-16b)$$

где S — как и раньше, площадь переходов.

Используя граничные условия (4-16), определяем коэффициенты A_1 и A_2 , входящие в выражение (4-15). После этого распределение концентрации можно привести к следующему виду:

$$p(x) = \frac{I_{\text{эп}}L}{qDS} \frac{\text{sh}\left(\frac{w-x}{L}\right)}{\text{ch}\left(\frac{w}{L}\right)} + p_0 \left[\left(e^{\frac{U_K}{\Psi_T}} - 1 \right) \frac{\text{ch}\left(\frac{x}{L}\right)}{\text{ch}\left(\frac{w}{L}\right)} + 1 \right]. \quad (4-17a)$$

При нормальном рабочем режиме ($U_K < 0$, $|U_K| \gg \Psi_T$), а также при обычном условии $w < L$ второй член в правой части (4-17a) не превышает $0,2 p_0$ и им можно пренебречь¹. Тогда

$$p(x) = \frac{I_{\text{эп}}L}{qDS} \frac{\text{sh}\left(\frac{w-x}{L}\right)}{\text{ch}\left(\frac{w}{L}\right)}. \quad (4-17b)$$

Если пренебречь рекомбинацией в базе, положив $L \rightarrow \infty$ (т. е. $w \ll L$), то

$$p(x) = \frac{I_{\text{эп}}w}{qDS} \left(1 - \frac{x}{w} \right). \quad (4-17b)$$

Как видим, стационарное распределение дырок в тонкой базе близко к линейному, на что уже обращалось внимание выше.

Дифференцируя (4-17b) по x , умножая обе части на $-qDS$ и полагая $x = w$, получаем для коллекторного тока выражение

$$I_K = \frac{I_{\text{эп}}}{\text{ch}(w/L)}.$$

Беря производную $dI_K/dI_{\text{эп}}$ или просто деля ток I_K на $I_{\text{эп}}$, находим коэффициент переноса:

$$\kappa = \frac{1}{\text{ch}\left(\frac{w}{L}\right)} = \text{sech}\left(\frac{w}{L}\right). \quad (4-18)$$

Формула (4-18) является одной из фундаментальных в теории транзисторов.

¹ Заметим, что для расчета коэффициента переноса второй член вообще несуществен (независимо от режима и толщины базы), поскольку в него не входит эмиттерный ток.

С увеличением толщины базы коэффициент переноса согласно (4-18) неограниченно уменьшается и при $w \gg L$ (практически при $w > 3L$) близок к нулю. Поэтому в транзисторе, как уже отмечалось, делают базу как можно более тонкой, так что всегда выполняется условие $w < L$. Тогда, разлагая $\operatorname{ch}(z)$ в ряд с точностью до трех первых членов, приходим к выражению

$$\kappa = \frac{1}{1 + \frac{1}{2} \left(\frac{w}{L}\right)^2 + \frac{1}{24} \left(\frac{w}{L}\right)^4}. \quad (4-19a)$$

Третий член знаменателя мал по сравнению со вторым, поэтому обычно записывают коэффициент переноса в виде

$$\kappa = \frac{1}{1 + \frac{1}{2} \left(\frac{w}{L}\right)^2}. \quad (4-19б)$$

Учитывая, что второй член знаменателя много меньше единицы, можно пользоваться еще одним приближением:

$$\kappa \approx 1 - \frac{1}{2} \left(\frac{w}{L}\right)^2. \quad (4-19в)$$

Например, при $w = (0,2 \div 0,3) L$ получаются значения $\kappa = 0,95 \div 0,98$.

Коэффициент передачи тока α согласно (4-13) получается путем умножения любого из выражений (4-18) или (4-19) на коэффициент инжекции. Так, из наиболее строгого выражения (4-18) следует:

$$\alpha = \gamma \operatorname{sech} \left(\frac{w}{L} \right), \quad (4-20)$$

а приближенные выражения (4-19б) и (4-19в) дают соответственно:

$$\alpha = \frac{\gamma}{1 + \frac{1}{2} \left(\frac{w}{L}\right)^2} \quad (4-21a)$$

и

$$\alpha \approx \gamma \left[1 - \frac{1}{2} \left(\frac{w}{L}\right)^2 \right]. \quad (4-21б)$$

Несмотря на ряд ограничений, общей тенденцией при разработке транзисторов является приближение коэффициента передачи α к единице, поскольку при этом, как увидим ниже, улучшаются усилительные свойства в схеме ОЭ.

В случае сплавных (бездрейфовых) транзисторов толщину базы w не удастся сделать менее 20—30 мкм, поэтому определяющую роль обычно играет коэффициент переноса. У дрейфовых транзисторов (см. § 4-12) значение w доходит до 1—2 мкм и главную роль обычно играет коэффициент инжекции.

Сопротивление эмиттерного перехода. Пусть коллекторное напряжение U_k имеет достаточно большую отрицательную величину: $|U_k| \gg \varphi_T$. Тогда, продифференцировав выражение (4-9), получим сопротивление эмиттерного перехода, которое нетрудно представить в следующей форме, аналогичной (2-38):

$$r_{\text{э}} = \frac{dU_{\text{э}}}{dI_{\text{э}}} = \frac{\varphi_T}{I_{\text{э}}}. \quad (4-22)$$

Как видим, это сопротивление обратно пропорционально току. При токе $I_{\text{э}} = 1$ мА оно составляет около 25 Ом. При меньших токах сопротивление $r_{\text{э}}$ увеличивается до сотен и тысяч ом.

Так, если $U_{\text{э}} = 0$, то

$$r_{\text{э}} \approx \frac{\varphi_T}{I'_{\text{э}0}},$$

а если $I_{\text{э}} = 0$, то

$$r_{\text{э}} \approx \frac{\varphi_T}{I'_{\text{э}0}(1 - \alpha_N)}.$$

Эти формулы легко получаются из выражения (4-7) при $|U_k| \gg \varphi_T$.

Сопротивление коллекторного перехода. Величина r_k , как уже известно, обусловлена эффектом модуляции толщины базы. Поэтому сначала оценим количественно этот эффект. Примем, что коллекторный переход полностью сосредоточен в базе; тогда

$$dw = -dl,$$

где l — ширина перехода, определяемая выражением (2-12).

Дифференцируя (2-12) по $|U|$ и полагая $|U| = U_k$, приходим к следующему результату:

$$dw = - \sqrt{\frac{\varepsilon_0 \varepsilon}{2qN_d}} \frac{dU_k}{\sqrt{U_k}}. \quad (4-23)$$

Как видно, эффект модуляции выражен тем сильнее, чем меньше напряжение U_k и чем больше удельное сопротивление базы, которое обратно пропорционально концентрации доноров N_d .

Теперь воспользуемся зависимостью (4-8), подставив в нее (4-21б):

$$I_k - I_{k0} = \alpha I_{\text{э}} = \gamma \left(1 - \frac{w^2}{2L^2} \right) I_{\text{э}}.$$

Дифференцируя I_k по w (при $I_{k0} = \text{const}$) и подставляя dw из соотношения (4-23), получаем следующее выражение для сопротивления коллекторного перехода:

$$r_k = \frac{dU_k}{dI_k} = \sqrt{\frac{2qN_d}{\varepsilon_0 \varepsilon}} \frac{L^2}{w} \frac{\sqrt{U_k}}{\gamma I_{\text{э}}}. \quad (4-24)$$

Пусть, например, $N_d = 10^{15}$ см⁻³; $L = 0,1$ мм; $w = 30$ мкм; $U_k = 5$ В; $I_{\text{э}} = 1$ мА; тогда для кремния ($\varepsilon = 12$) $r_k \approx 4,2$ МОм.

Заметим, что сопротивление r_k , как и сопротивление r_a , обратно пропорционально току эмиттера. Однако в отличие от r_a сопротивление r_k зависит еще от ряда параметров, в том числе от напряжения U_k .

Коэффициент обратной связи по напряжению. Из рис. 4-8, а легко видеть, что приращение граничной концентрации дырок определяется простым соотношением

$$\frac{dp(0)}{p(0)} = \frac{d\omega}{\omega}.$$

Далее из формулы (2-13а), беря производную по U и полагая $U = U_a$, находим:

$$\frac{dp(0)}{p(0)} = \frac{dU_a}{\varphi_T}.$$

Приравнивая правые части обоих соотношений и подставляя $d\omega$ из (4-23), нетрудно получить:

$$\mu_{\text{ЭК}} = \frac{dU_a}{dU_k} = - \sqrt{\frac{\epsilon_0 \epsilon}{2qN_d}} \frac{\varphi_T}{\omega \sqrt{U_k}}. \quad (4-25)$$

При значениях параметров, использованных для оценки r_k по формуле (4-24), $\mu_{\text{ЭК}} \approx -1,1 \cdot 10^{-4}$. Знак минус говорит о том, что увеличение коллекторного напряжения (по модулю) уменьшает эмиттерное напряжение. Малая величина $\mu_{\text{ЭК}}$ означает слабое смещение кривых на рис. 4-11, б при изменении параметра U_k . Например, если $\Delta U_k = 3$ В, то в нашем примере $|\Delta U_a| \approx 0,3$ мВ.

Перемножая правые части (4-24), (4-25) и учитывая выражения (4-22) и (4-21б), легко получить соотношение между внутренними параметрами транзистора:

$$2\mu_{\text{ЭК}} r_k (\gamma - \alpha) = r_a. \quad (4-26)$$

Иногда обратную связь по напряжению отражают не эквивалентным генератором э. д. с. $\mu_{\text{ЭК}} U_k$, как сделано на рис. 4-13, а дополнительным сопротивлением в цепи базы, включенным последовательно с r_b (см., например, [58]). Значение этого дополнительного сопротивления, называемого *диффузионным сопротивлением базы* ($r_{b.d}$), находится из условия

$$\mu_{\text{ЭК}} U_k \approx \frac{r_{b.d}}{r_{b.d} + r_k} U_k,$$

в котором правая часть представляет собой часть напряжения U_k , передаваемую в цепь эмиттера через делитель $r_{b.d} - r_k$. Полагая $r_{b.d} \ll r_k$, $\gamma = 1$ и учитывая (4-26), получаем:

$$r_{b.d} \approx \mu_{\text{ЭК}} r_k = \frac{r_a}{2(1 - \alpha)}.$$

Можно показать, что, вводя диффузионное сопротивление базы, необходимо одновременно заменить на эквивалентной схеме сопротивление r_a на $r_a/2$ для того, чтобы не изменилось входное сопротивление. Это обстоятельство, а также тот факт, что на повышенных частотах сопротивление $r_{b.d}$ становится комплексным (так как при его расчете нужно вместо r_k использовать Z_k), делают данную экви-

валентную схему несколько искусственной мы не будем ее применять, тем более, что обратную связь по напряжению вообще приходится учитывать сравнительно редко.

Объемное сопротивление базы. Идеализируем структуру сплавного транзистора так, как показано на рис. 4-14. Как видим, базовый ток (если считать, что он протекает от центра базы к периферии) встречает различные сопротивления на трех разных участках. Первый — внутренний — участок (активная область базы) является диском с толщиной w_1 и радиусом R_1 . Второй и третий участки (промежуточная и пассивная области базы) являются кольцами с толщинами w_2 и w_3 и внешними радиусами соответственно R_2 и R_3 .

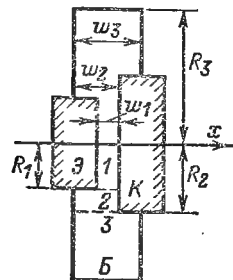


Рис. 4-14. Идеализированная структура сплавного транзистора.

Сопротивления колец в радиальном направлении выражаются формулами

$$r_{62} = \frac{\rho_6}{2\pi w_2} \ln \frac{R_2}{R_1};$$

$$r_{63} = \frac{\rho_6}{2\pi w_3} \ln \frac{R_3}{R_2},$$

где ρ_6 — удельное сопротивление базы.

Сопротивление диска, у которого внутренний диаметр равен нулю, нельзя рассчитать по аналогичной формуле. Поэтому для оценки сопротивления первого участка воспользуемся следующим приемом: найдем падение напряжения вдоль радиуса активной области базы и поделим это напряжение на ток базы.

Напомним, что стационарный ток базы I_6 обусловлен рекомбинацией (см. с. 179) и потому пропорционален всему объему активной области ($\pi R_1^2 w_1$). Если внутри активной области выделить цилиндрический объем с радиусом $R < R_1$, то ток $I_6(R)$, вытекающий из этого цилиндра, будет пропорционален объему ($\pi R^2 w_1$). Следовательно,

$$I_6(R) = I_6 \frac{R^2}{R_1^2}.$$

Сопротивление элементарного кольца с внутренним радиусом R и внешним радиусом $R + dR$ будет

$$dr_{61} = \rho_6 \frac{dR}{2\pi R w_1}.$$

Интегрируя элементарное падение напряжения $I_6(R) dr_{61}$ в пределах от 0 до R_1 и деля на ток I_6 , получаем искомое сопротивление¹:

$$r_{61} = \frac{0,5\rho_6}{2\pi w_1}.$$

Суммируя полученные выражения, запишем полное сопротивление базы:

$$r_6 = \frac{\rho_6}{2\pi} \left(\frac{0,5}{w_1} + \frac{1}{w_2} \ln \frac{R_2}{R_1} + \frac{1}{w_3} \ln \frac{R_3}{R_2} \right). \quad (4-27)$$

Пусть, например, $\rho_6 = 5$ Ом·см; $w_1 = 40$ мкм; $w_2 = 5 w_1$; $w_3 = 9 w_1$; $R_2 = 1,5 R_1$; $R_3 = 5 R_1$; тогда $r_6 \approx 140$ Ом. Практи-

¹ В [59] вместо коэффициента 0,5 получено значение 0,25.

чески у всех транзисторов главную роль играет внутренний участок (активная область базы), имеющий наименьшую толщину; его сопротивление в нашем примере составляет примерно 100 Ом.

Заметим, что выражение (4-27) и вывод относительно главной роли сопротивления $r_{б1}$ действительны лишь при достаточно малых токах, когда можно пренебречь модуляцией сопротивления базы инжектированными носителями (см. § 2-8). При высоких уровнях инжекции удельное сопротивление активной базы заметно уменьшается, так что результирующее значение $r_б$ становится близким к сопротивлению пассивной области.

Кроме того, следует иметь в виду, что сплавные транзисторы, применительно к которым проведены расчеты, в настоящее время имеют в основном историческое значение. Главную роль сейчас играют планарные транзисторы и их варианты. Структура планарного транзистора (см. рис. 4-53) существенно отличается от структуры сплавного; поэтому расчет сопротивления базы оказывается сложнее; оно выражается иной формулой и иным оказывается соотношение между вкладом активной и пассивной частей. Однако в основе расчета, как и в данном разделе, лежат чисто геометрические соображения.

Тепловой ток коллектора. Чтобы найти ток $I_{к0}$, можно воспользоваться распределением (4-17а), положив $I_{эp} = 0$ и $|U_{к}| \geq \varphi_{г}$. Тогда ¹

$$p(x) = -p_0 \frac{\operatorname{ch}\left(\frac{x}{L}\right)}{\operatorname{ch}\left(\frac{w}{L}\right)} + p_0. \quad (4-28)$$

Дифференцируя (4-28) по x , полагая $x = w$ и умножая на $-qDS$, получаем искомое значение теплового тока:

$$I_{к0} \approx \frac{qDS p_0}{L} \operatorname{th}\left(\frac{w}{L}\right). \quad (4-29a)$$

Учитывая соотношение $w \ll L$, при котором $\operatorname{th} \frac{w}{L} \approx w/L$, получаем:

$$I_{к0} \approx \frac{qDS p_0}{L} \frac{w}{L}. \quad (4-29б)$$

В формулах (4-29) не учтена электронная составляющая тока, которая при условии $r_{к} \ll r_б$ не существенна.

Легко показать, что в рассматриваемом режиме напряжение на эмиттерном переходе имеет небольшую отрицательную вели-

¹ Согласно (4-28) концентрация дырок близка к нулю во всей базе: даже на эмиттерной границе, где концентрация максимальна, ее значение составляет:

$$p(0) = p_0 \left[1 - \operatorname{ch}^{-1}\left(\frac{w}{L}\right) \right] \approx p_0 \frac{1}{2} \left(\frac{w}{L} \right)^2.$$

При $w/L \leq 0,3$ получается $p(0) \leq 0,05 p_0$.

чину (φ_s на рис. 4-5, б). Подставляя $I_s = 0$ и $|U_k| \gg \varphi_T$ в общее выражение (4-7), получаем:

$$\varphi_s = \varphi_T \ln(1 - \alpha_N).$$

Если $\alpha_N = 0,98$, то $\varphi_s \approx -4\varphi_T \approx -0,1$ В.

4-5. ДИНАМИЧЕСКИЕ ПАРАМЕТРЫ ТРАНЗИСТОРА

Барьерные емкости. Эти емкости не имеют особой специфики по сравнению с емкостями p - n перехода (диода) и могут быть рассчитаны по формулам (2-82). При этом, как уже отмечалось, формула (2-82б) оказывается неточной при сколько-нибудь значительных прямых смещениях. Поэтому емкость C_s в рабочем режиме обычно оценивают из эмпирического соотношения $C_s = (1,5-1,7)C_{s0}$, где C_{s0} — емкость при $U_s = 0$.

Относительно емкости C_k нужно заметить, что она согласно [60] представляет собой емкость коллектора, соответствующую a к т и в н о й области базы, т. е. вычисляется для площади $S_k = S_s$. Емкости коллектора, соответствующие промежуточной и пассивной областям базы, могут считаться соединенными параллельно и подключенными между внешними электродами коллектор — база. Эта межэлектродная емкость часто сравнима со значением C_k , однако при расчетах ее можно относить к нагрузке транзистора.

Основной интерес при анализе переходных и частотных характеристик транзистора представляет коэффициент передачи тока α . Поскольку согласно (4-13) коэффициент α состоит из двух множителей — коэффициента инжекции и коэффициента переноса, рассмотрим эти множители отдельно.

Коэффициент инжекции. Напомним, что ток, заряжающий барьерную емкость C_s , образуется основными носителями (см. с. 148) и поэтому не передается в цепь коллектора. Значит, увеличение емкостного тока на высоких частотах равносильно уменьшению коэффициента инжекции¹:

$$\dot{\gamma} = \frac{\gamma}{1 + r_s/X_c} = \frac{\gamma}{1 + j \frac{\omega}{\omega_y}}, \quad (4-30a)$$

где $X_c = 1/j\omega C_s$.

¹ В определении статического коэффициента инжекции γ (с. 118) полный ток перехода $I(0)$ был равен сумме $I_p(0) + I_n(0)$. В данном случае, применительно к динамическому коэффициенту инжекции, $I(0) = I_p(0) + I_n(0) + I_c(0)$. Отсюда следует:

$$\dot{\gamma} = \frac{I_p(0)}{I(0)} = \frac{\gamma}{1 + \frac{I_c(0)}{I_p(0) + I_n(0)}} = \frac{\gamma}{1 + I_c(0)/I_{p-n}(0)}.$$

Поскольку барьерная емкость и p - n переход включены параллельно, отношение их токов обратно пропорционально отношению их сопротивлений (X_c и r_s); после подстановки получается выражение (4-30a).

Соответственно изображение коэффициента инжекции имеет вид:

$$\gamma(s) = \frac{\gamma}{1 + s\tau_\gamma}. \quad (4-306)$$

В формулах (4-30) граничная частота ω_γ (на уровне 0,7) и постоянная времени τ_γ определяются выражением

$$\tau_\gamma = \frac{1}{\omega_\gamma} = r_s C_s. \quad (4-31)$$

Как видим, частотные и переходные свойства коэффициента γ улучшаются с увеличением рабочего тока, так как при этом уменьшается сопротивление r_s . При токе 2 мА (когда $r_s \approx 12$ Ом) и при $C_s = 25$ пФ граничная частота f_γ составляет примерно 500 МГц.

У дрейфовых транзисторов, характерных большими рабочими токами и меньшими емкостями, частота f_γ достигает тысяч мегагерц.

Коэффициент переноса. В данном разделе рассматриваются только операторные изображения и временные параметры коэффициента переноса, а соответствующие переходные и частотные характеристики анализируются в следующем разделе.

Для того чтобы найти изображение $\kappa(s)$, нужно решить полное уравнение диффузии (1-79а). Поскольку в начальном состоянии $\Delta p = 0$ во всей базе, изображением производной $\partial p / \partial t$ будет просто $s\Delta p$. Тогда в операторном виде уравнение (1-79а) будет таким же, как (2-89). Запишем его для полной концентрации, подставляя $\Delta p = p - p_0$:

$$\frac{d^2 p}{dx^2} - \frac{p}{L^2(s)} = - \frac{p_0}{L^2(s)}, \quad (4-32)$$

где $L(s)$ — операторная диффузионная длина (1-122).

Как видим, уравнение (4-32) отличается от (4-14) только заменой L на $L(s)$. Значит, при тех же граничных условиях (4-16) искомая функция $\kappa(s)$ будет определяться выражением (4-18) с заменой L на $L(s)$:

$$\kappa(s) = \operatorname{sech} \left[\frac{w}{L(s)} \right] = \operatorname{sech} \left(\frac{w}{L} \sqrt{1 + s\tau} \right). \quad (4-33)$$

Оригиналом этого изображения является сложный ряд, неудобный для практического использования. Поэтому широкое распространение имеют различные аппроксимации.

Наиболее простая аппроксимация (*1-е приближение*) получается при разложении $\operatorname{ch}(z)$ в ряд с точностью до двух первых членов [см. (4-19)]. В этом случае изображение $\kappa(s)$ легко приводится к виду

$$\kappa(s) = \frac{\kappa}{1 + s\tau_\kappa}, \quad (4-34)$$

откуда ясно, что оригиналом является простейшая экспонента. Ее постоянная времени выражается следующим образом:

$$\tau_k = \kappa \frac{\omega^2}{2L^2} \tau.$$

Подстановка $L^2 = D\tau$ приводит к выводу, что постоянная времени τ_k практически равна времени диффузии (2-88):

$$\tau_k = \kappa \frac{\omega^2}{2D} = \kappa \tau_D. \quad (4-35a)$$

Если же величину $\omega^2/2L^2$ выразить через κ с помощью (4-19б), получается связь постоянной времени τ_k с временем жизни в виде соотношения

$$\tau_k = (1 - \kappa) \tau. \quad (4-35б)$$

Формула (4-35a) говорит о том, что частотные и переходные свойства транзистора улучшаются прежде всего с уменьшением толщины базы. Кроме того, важную роль играет значение коэффициента диффузии, т. е. подвижности неосновных носителей. Последнее обстоятельство поясняет интерес к таким материалам, у которых подвижность носителей больше, чем у германия и кремния. Ясно также, что кремниевые транзисторы при прочих равных условиях уступают германиевым по динамическим свойствам из-за меньшей подвижности носителей.

Тот факт, что при разложении в ряд функции (4-33) мы пренебрегли высшими степенями оператора s , означает заведомую неточность в начальной части переходной характеристики. Точная функция $\kappa(t)$ (см., например, [14]) характеризуется несколько меньшей постоянной времени τ_k по сравнению с (4-35) и наличием небольшой задержки начала фронта (диффузионная задержка).

Чтобы уточнить изображение $\kappa(s)$, воспользуемся снова функцией (4-33), но, разлагая $\operatorname{sh}(z)$ в ряд, оставим не два, а три первых члена (2-е приближение). Тогда после некоторых преобразований получаем:

$$\kappa(s) = \frac{\kappa}{1 + b_1 s + b_2 s^2}, \quad (4-36a)$$

где коэффициенты b после подстановки $L^2 = D\tau$ имеют вид:

$$b_1 = \kappa \tau_D; \quad b_2 = \frac{1}{6} \kappa \tau_D^2.$$

Оригиналом изображения (4-36a) является функция, содержащая две экспоненты с разными постоянными времени. Начальная производная этой функции равна нулю, что соответствует наличию задержки фронта.

Последующие выводы будут более наглядными, если изображение (4-36a) записать в другой форме:

$$\kappa(s) = \frac{\kappa}{(1 + s\tau_1)(1 + s\tau_2)}, \quad (4-36б)$$

где τ_1 и τ_2 — постоянные времени тех двух экспонент, о которых сказано выше.

Приравнявая знаменатели (4-36а) и (4-36б), получаем: $b_1 = \tau_1 + \tau_2$ и $b_2 = \tau_1 \tau_2$. Решая эти уравнения относительно постоянных времени и учитывая значения коэффициентов b , получаем следующие выражения:

$$\tau_1 = \frac{1}{2} \kappa t_D (1 + \sqrt{1 - 2/3\kappa}); \quad (4-37a)$$

$$\tau_2 = \frac{1}{2} \kappa t_D (1 - \sqrt{1 - 2/3\kappa}). \quad (4-37б)$$

В реальных условиях (когда $\kappa \approx 1$) постоянная времени τ_1 в несколько раз превышает постоянную времени τ_2 .

Из общей теории переходных процессов следует, что при таком соотношении величина τ_1 определяет длительность, а величина τ_2 — задержку экспоненциального фронта¹. Соответственно можно считать

$$\tau_{\kappa} = \tau_1; \quad (4-38a)$$

$$t_{3\kappa} = \tau_2. \quad (4-38б)$$

Если положить $\kappa = 1$, то согласно (4-37) и (4-38)

$$\tau_{\kappa} \approx 0,8 t_D; \quad (4-39a)$$

$$t_{3\kappa} \approx 0,2 t_D. \quad (4-39б)$$

Таким образом, изображения (4-36) можно заменить а п п р о к с и м а ц и е й

$$\kappa(s) = \frac{\kappa e^{-st_{3\kappa}}}{1 + s\tau_{\kappa}}, \quad (4-40)$$

где время задержки $t_{3\kappa}$ фигурирует в явном виде. Такая аппроксимация по сравнению с (4-36) нагляднее и удобнее при использова-

¹ Известно (см. [62, с. 89]), что если коэффициент передачи имеет изображение

$$\frac{1 + a_1 s + a_2 s^2 + \dots}{1 + b_1 s + b_2 s^2 + \dots},$$

то время нарастания фронта (на уровне 0,1—0,9) выражается следующим образом:

$$t_H = 2,2 \sqrt{b_1^2 - a_1^2 + 2(a_2 - b_2)}.$$

В случае (4-36а) имеем $a_1 = a_2 = 0$. Подставляя приведенные в тексте значения b_1 и b_2 и полагая $\tau_1 \gg \tau_2$, получаем:

$$t_H = 2,2 \tau_1 \sqrt{1 + (\tau_2/\tau_1)^2} \approx 2,2 \tau_1.$$

Как видим, формирование фронта происходит почти так же, как если бы в изображении (4-36б) постоянная времени τ_2 равнялась нулю. Формально такой вывод дает основание считать $s\tau_2 \ll 1$. Тогда сумму $1 + s\tau_2$ можно рассматривать как первые два члена разложения функции $e^{s\tau_2}$. Эта функция, будучи перенесена в числитель в виде $e^{-s\tau_2}$, означает (согласно законам операционного исчисления) сдвиг отсчета времени и переходной характеристики на величину τ_2 .

нии. Часто выражение (4-40) записывают в виде

$$\kappa(s) = \frac{\kappa e^{-s(c\tau_\kappa)}}{1 + s\tau_\kappa}, \quad (4-41)$$

где $c = t_{3\kappa}/\tau_\kappa \approx 0,25$ — так называемый *фактор задержки*.

Коэффициент передачи тока. В общем случае динамические характеристики коэффициента α определяются обоими параметрами γ и κ . Однако, если $\tau_\gamma \ll \tau_\kappa$, то переходные и частотные характеристики коэффициента α определяются только коэффициентом переноса, т. е. действительны формулы предыдущего раздела. Начнем анализ с этого случая.

Умножая обе части (4-33) на γ , получаем изображение

$$\alpha(s) = \gamma \operatorname{sech}\left(\frac{\omega}{L} \sqrt{1 + s\tau}\right), \quad (4-42a)$$

которому соответствует частотная характеристика

$$\dot{\alpha} = \gamma \operatorname{sech}\left(\frac{\omega}{L} \sqrt{1 + j\omega\tau}\right). \quad (4-42b)$$

Обе эти функции не имеют практического распространения. Используя 1-е приближение (4-34), получаем:

$$\alpha(s) = \frac{\alpha}{1 + s\tau_\alpha}, \quad (4-43)$$

где

$$\tau_\alpha = \tau_\kappa = \kappa t_D \approx \frac{\omega^2}{2D}. \quad (4-44)$$

Оригиналом изображения (4-43) является простейшая экспоненциальная функция

$$\alpha(t) = \alpha(1 - e^{-t/\tau_\alpha}), \quad (4-45)$$

у которой начальная производная имеет конечное значение α/τ_α , т. е. задержка фронта отсутствует.

Частотная характеристика, соответствующая изображению (4-43), получается заменой оператора s на $j\omega$ и имеет вид:

$$\dot{\alpha} = \frac{\alpha}{1 + j \frac{\omega}{\omega_\alpha}}. \quad (4-46)$$

Амплитудно-частотная и фазо-частотная характеристики легко получаются из (4-46):

$$|\dot{\alpha}| = \frac{\alpha}{\sqrt{1 + (\omega/\omega_\alpha)^2}}; \quad (4-47a)$$

$$\varphi = -\arctg(\omega/\omega_\alpha). \quad (4-47b)$$

Граничная частота (на уровне 0,7) равна:

$$\omega_\alpha = \frac{1}{\tau_\alpha} \approx \frac{2D}{\omega^2}. \quad (4-48)$$

Фазовый сдвиг на граничной частоте равен: $\varphi_\alpha = -45^\circ$.

Изображение (4-43) и связанные с ним функции (4-45) — (4-47) широко используются при анализе схем, если задержка фронта и сдвиг фазы не играют первоочередной роли.

Динамические характеристики, основанные на 2-м приближении (4-36), оказываются гораздо более точными. Так, переходная характеристика имеет нулевую начальную производную, что отражает задержку фронта. Граничная частота получается на 20% выше, чем (4-48). Фазовый сдвиг на граничной частоте равен: $\varphi_\alpha \approx -53^\circ$.

Однако, несмотря на большую точность, 2-е приближение неудобно при практических расчетах. Поэтому, если 1-е приближение по тем или иным причинам неприемлемо, используют аппроксимацию (4-40), которая применительно к коэффициенту передачи тока имеет вид:

$$\alpha(s) = \frac{\alpha e^{-st_{sa}}}{1 + st_{sa}}. \quad (4-49)$$

Здесь согласно (4-39)

$$\tau_\alpha = \tau_\kappa = 0,8t_D; \quad (4-50a)$$

$$t_{sa} = t_{sn} = 0,2t_D. \quad (4-50b)$$

Оригиналом изображения (4-49) является обычная экспоненциальная характеристика, но сдвинутая относительно момента $t = 0$ на время t_{sa} :

$$\alpha(t) = 0, \quad t \leq t_{sa}; \quad (4-51a)$$

$$\alpha(t) = \alpha(1 - e^{-t/\tau_\alpha}), \quad t \geq t_{sa}. \quad (4-51b)$$

Частотная характеристика получается из (4-49) в виде

$$\dot{\alpha} = \frac{\alpha e^{-j\omega t_{sa}}}{1 + j \frac{\omega}{\omega_\alpha}}. \quad (4-52)$$

Амплитудно-частотная характеристика по форме совпадает с выражением (4-47a):

$$|\dot{\alpha}| = \frac{\alpha}{\sqrt{1 + (\omega/\omega_\alpha)^2}}, \quad (4-53)$$

но граничная частота количественно отличается от (4-48) в связи с различием выражений (4-44) и (4-50a), а именно:

$$\omega_\alpha = \frac{1}{\tau_\alpha} = \frac{2,5D}{\tau^2}. \quad (4-54)$$

Фазо-частотная характеристика принципиально отличается от соответствующих характеристик 1-го и 2-го приближений:

$$\varphi = -\omega t_{sa} - \arctg\left(\frac{\omega}{\omega_\alpha}\right). \quad (4-55)$$

Отличие состоит в том, что в данном случае фаза неограниченно возрастает с увеличением частоты, тогда как в 1-м приближении $\varphi(\infty) = -90^\circ$, а во 2-м приближении $\varphi(\infty) = -180^\circ$. Что касается сдвига фазы на граничной частоте, то в данном случае $\varphi_\alpha \approx -59^\circ$; это значение почти совпадает с точным значением -57° , получаемым из выражения (4-42б).

На рис. 4-15 показаны оригинал функции (4-42а), а также оригинал 1-го приближения (4-43). Оригинал 2-го приближения практически сливается со строгим решением, а оригинал аппроксимации (4-49) отличается от строгого решения только в начальной части, поскольку функция (4-51) начинается из точки $t = 0,2 t_D$.

На рис. 4-16 показаны точная и приближенные частотные характеристики, соответствующие функциям (4-42б), (4-46) и 2-му приближению. Характеристики, соответствующие функциям (4-52), практически совпадают с точными кривыми.

Перейдем теперь к оценке влияния постоянной времени τ_γ . Учет этой постоянной времени сводится к перемножению изображений $\kappa(s)$ и $\gamma(s)$. При этом степень операторного полинома в изображении $\alpha(s)$ повышается и в соответствующем оригинале появляется дополнительная экспонента с постоянной времени τ_γ . Такие функции, как уже отмечалось, на практике неудобны. Поэтому целесообразно путем приближений сохранить изображение типа (4-49). Обычно это удается сделать без большой погрешности.

Примем для коэффициента переноса функцию (4-40). Умножая ее на (4-30б), получаем изображение $\alpha(s)$ в общем виде:

$$\alpha(s) = \frac{\alpha e^{-s t_{3\kappa}}}{(1 + s \tau_\gamma)(1 + s \tau_\kappa)}. \quad (4-56)$$

Пусть сначала $\tau_\gamma \ll \tau_\kappa$. Тогда согласно общим свойствам такого рода изображений (см. сноску на с. 199) можно считать τ_κ постоянной времени, а τ_γ — дополнительной задержкой. Иначе говоря, в случае $\tau_\gamma \ll \tau_\kappa$ действительно изображение (4-49) с параметрами ¹

$$\tau_\alpha \approx \tau_\kappa = 0,8 t_D; \quad (4-57a)$$

$$t_{3\alpha} \approx t_{3\kappa} + \tau_\gamma = 0,2 t_D + \tau_\gamma. \quad (4-57б)$$

¹ При условии $\tau_\gamma \ll 0,2 t_D$ роль постоянной времени τ_γ согласно (4-57б) становится вообще пренебрежимо малой. В этом случае на эквивалентной схеме транзистора (рис. 4-13) можно не учитывать наличия емкости C_β , что характерно для низкочастотных транзисторов.

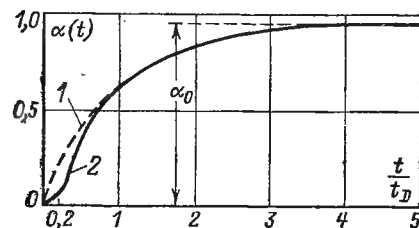


Рис. 4-15. Переходные характеристики коэффициента передачи.
1 — первое приближение; 2 — строгое решение.

Пусть теперь, наоборот, $\tau_y \gg \tau_k$. Тогда из тех же соображений можно считать τ_y постоянной времени, а τ_k дополнительной задержкой. Значит, и в этом случае действительно изображение (4-49), но с параметрами

$$\tau_\alpha \approx \tau_y = r_3 C_3; \quad (4-58a)$$

$$t_{3\alpha} \approx t_{3k} + \tau_k \approx t_D. \quad (4-58b)$$

Менее типичен случай, когда постоянные времени τ_y и τ_k имеют сравнимые величины, т. е. различаются не более чем в 1,5—2 раза. При этом изображение (4-56) можно приближенно записать в виде

$$\alpha(s) \approx \frac{\alpha e^{-st_{3k}}}{(1 + s\tau_{yк})^2}, \quad (4-59)$$

где постоянная времени $\tau_{yк}$ близка к среднеарифметическому обоих компонентов:

$$\tau_{yк} \approx \frac{1}{2} (\tau_y + \tau_k).$$

Функция (4-59) практически столь же неудобна, как и (4-56). Однако с некоторой погрешностью ее можно заменить функцией (4-49), если в последней положить

$\tau_\alpha = \sqrt{2} \tau_{yк}$. При этом времена нарастания, вычисленные по изображениям (4-59) и (4-49), будут одинаковыми.

На рис. 4-17 влияние постоянной времени τ_y подытожено

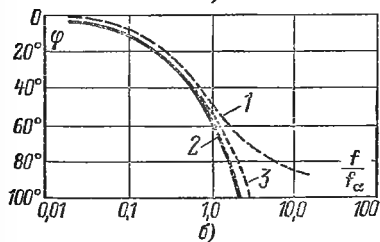
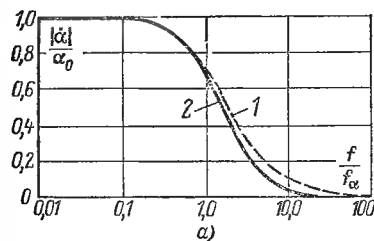


Рис. 4-16. Частотные характеристики коэффициента передачи.

а — амплитудно-частотные; б — фазо-частотные.

Кривые означают: 1 — первое приближение; 2 — строгое решение; 3 — второе приближение.

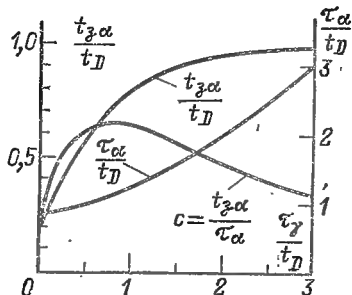


Рис. 4-17. Зависимость динамических параметров коэффициента передачи тока от постоянной времени эмиттерной цепи.

В виде кривых, которые позволяют при любом значении τ_y найти соответствующие значения τ_α и $t_{3\alpha}$.

Часто в соответствии с (4-41) изображение $\alpha(s)$ записывают в виде

$$\alpha(s) = \frac{\alpha e^{-s(c\tau_\alpha)}}{1 + s\tau_\alpha}, \quad (4-60)$$

где фактор задержки $c = t_{3\alpha}/\tau_{\alpha}$. Как видно из рис. 4-17, фактор задержки может иметь значения от 0,25 до 0,66 в зависимости от значения τ_y . Если $\tau_y = 0$ (точнее, если $\tau_y \ll 0,2 t_D$), то обычно принимают $c = 0,25$.

Частотные характеристики коэффициента α с учетом τ_y можно описывать единым выражением (4-52), заимствуя параметры $t_{3\alpha}$ и τ_{α} из рис. 4-17. Иногда выражение (4-52) записывают в форме, аналогичной (4-60):

$$\dot{\alpha} = \frac{\alpha e^{-jc(\omega/\omega_{\alpha})}}{1 + j \frac{\omega}{\omega_{\alpha}}}, \tag{4-61}$$

где при достаточно малых значениях τ_y полагают $c = 0,25$.

Общий важный вывод состоит в том, что влияние эмиттерной емкости — параметра, не связанного с процессами в базе, — учитывается путем изменения параметров коэффициента переноса. Тем самым удастся избежать усложнения аналитических выражений, характеризующих переходный процесс передачи тока.

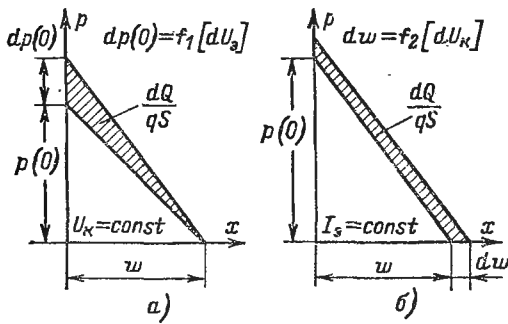


Рис. 4-18. Изменения заряда в базе, обуславливающие диффузионные емкости эмиттера (а) и коллектора (б).

Диффузионные емкости. Понятие диффузионной емкости было введено при рассмотрении диодов (см. § 2-9) как параметр, характеризующий зависимость приращения заряда в

базе от приращения напряжения на переходе. Соответственно в транзисторах различают две диффузионные емкости — эмиттерную и коллекторную.

Будем считать распределение дырок в базе линейным (рис. 4-18), что имеет место при условии $w < L$. Тогда неравновесный заряд дырок будет выражаться формулой (2-87а), если ток I заменить на I_s :

$$\Delta Q = I_s \frac{w^2}{2D} = I_s t_D. \tag{4-62}$$

Дифференцируя (4-62) по напряжению U_s и учитывая (4-22), получаем диффузионную емкость эмиттера:

$$C_{э.д} = \frac{t_D}{r_s} = I_s \frac{t_D}{\varphi_T}. \tag{4-63а}$$

Например, если $t_D = 0,1$ мкс и $r_s = 25$ Ом, то $C_{э.д} = 4000$ пФ. Перезарядка емкости $C_{э.д}$ связана с распространением инжектированных носителей вдоль базы, т. е. отражает тот же процесс,

что и переходная характеристика коэффициента η . Иначе говоря, диффузионная емкость эмиттера заряжается не основными носителями и, следовательно, наличие этой емкости никак не отражается на коэффициенте инжекции. Поэтому емкость $C_{э.д}$ ни в коем случае не должна входить в выражение (4-31). Сумма $C_{э} + C_{э.д}$ характеризует лишь инерционность напряжения на эмиттерном переходе при заданной ступеньке тока $I_{э}$. Во избежание недоразумений емкости $C_{э}$ и $C_{э.д}$ рисуются на эквивалентной схеме раздельно.

Более строгий анализ показывает, что напряжение на эмиттерном переходе при скачке тока $I_{э}$ меняется не по экспоненциальному закону, т. е. постоянная времени $\tau_{э} (C_{э} + C_{э.д})$ неточно отражает переходный процесс. Однако, чтобы не усложнять расчеты, переходный процесс $u_{э}(t)$ все же считают экспоненциальным, но постоянную времени несколько уменьшают. Тогда диффузионная емкость эмиттера равна [58]:

$$C_{э.д} = \frac{2}{3} \frac{t_D}{r_{э}}. \quad (4-63)$$

Дифференцируя (4-62) по напряжению $U_{к}$ (при $I_{э} = \text{const}$), получаем диффузионную емкость коллектора. Используя выражения (4-23) и (4-24), можно привести ее к виду

$$C_{к.д} = V \sqrt{\frac{\epsilon_0 \epsilon}{2qN_d}} \frac{I_{э} \omega}{D V \ddot{U}_{к}} = \frac{\tau}{r_{к}}. \quad (4-64)$$

Например, если $\tau = 2$ мкс и $r_{к} = 1$ МОм, то $C_{к.д} = 2$ пФ.

Как видим, емкость $C_{к.д}$ несравненно меньше, чем $C_{э.д}$. Это объясняется разным механизмом влияния напряжений $U_{э}$ и $U_{к}$ на заряд в базе: приращение $\Delta U_{э}$ влияет на заряд непосредственно, меняя количество инжектируемых носителей, тогда как приращение $\Delta U_{к}$ влияет на заряд косвенно благодаря модуляции толщины базы (эффект Эрли).

Диффузионная емкость коллектора $C_{к.д}$ играет относительно малую роль, поскольку ее значение не превышает значения барьерной емкости $C_{к}$. Сумма $C_{к} + C_{к.д}$ определяет инерционность напряжения на коллекторном переходе по отношению к скачкам коллекторного тока. В остальных случаях эти емкости тоже нельзя суммировать, так как они перезаряжаются носителями разных типов: барьерная емкость — основными, а диффузионная — неосновными. В частности, инерционность обратной связи по напряжению обусловлена только емкостью $C_{к.д}$. Поскольку коэффициент $\mu_{э\kappa}$ очень мал (см. § 4-4), его инерционность незначительна, и поэтому ниже понятие диффузионной емкости коллектора не используется.

Диффузионная емкость эмиттера не будет специально изображаться на схемах; она должна учитываться лишь тогда, когда инерционность эмиттерного напряжения имеет самостоятельное значение.

Постоянная времени базы. До сих пор сопротивление базы по существу считалось отсутствующим, так как внешнее напряже-

ние U_k считалось приложенным непосредственно к коллекторному переходу.

С учетом сопротивления r_6 (см. рис. 4-13) ток генератора αI_9 не полностью идет во внешнюю цепь: часть тока ответвляется во внутреннюю цепь r_k, C_k . Таким образом, в реальном транзисторе при заданном токе I_9 ток I_k зависит не только от коэффициента α , но и от соотношения сопротивлений коллекторного перехода и базы. В области высоких частот и малых времен активной составляющей r_k можно пренебречь, так что токораспределение в этом случае обусловлено сопротивлениями r_6 и $X_c(s) = 1/sC_k$. Очевидно, что в е ш н и й ток коллектора будет связан с током эмиттера соотношением

$$I_k = \alpha(s) I_9 \frac{X_c(s)}{X_c(s) + r_6} = I_9 \frac{\alpha(s)}{1 + sr_6}, \quad (4-65)$$

где

$$\tau_6 = r_6 C_k. \quad (4-66)$$

Как видим, если принять для $\alpha(s)$ изображение (4-49), то выражение (4-65) будет иметь форму (4-56). Следовательно, выводы, которые были сделаны выше относительно влияния постоянной времени τ_v , остаются в силе и для данного случая применительно к постоянной времени τ_6 . А именно, поскольку практически у всех транзисторов (в том числе высокочастотных) выполняется условие $\tau_6 < 0,5 \tau_a$, то *постоянная времени τ_6 почти всегда играет роль дополнительной задержки фронта*, т. е. суммируется с временем задержки t_{3a} .

Однако цепочка r_6, C_k не только вносит поправку в динамические параметры коэффициента α . Более важное значение имеет тот факт, что эта цепочка обуславливает *о б р а т н у ю с в я з ь*, так как *через емкость C_k и сопротивление r_6 часть коллекторного напряжения передается на вход транзистора*. Такая обратная связь может быть существенной при небольшом сопротивлении источника тока I_9 . С ростом частоты обратная связь усиливается и на граничной частоте $\omega_6 = 1/\tau_6$ коэффициент передачи цепочки r_6, C_k составляет 0,7. На низких частотах ($\omega \ll \omega_6$) обратная связь пренебрежимо мала.

Инверсные параметры. Расчет инверсных параметров осложняется тем, что инжекция со стороны коллектора означает распространение дырок не только в активной области базы, но и в промежуточной области, т. е. в кольце 2 (рис. 4-14). Обычно площади эмиттера и коллектора различаются в 2—3 раза и более, так что роль промежуточной области оказывается весьма существенной. Не пытаясь провести сложный анализ неоднородной модели, рассмотрим особенности инверсных параметров, руководствуясь качественными соображениями.

Очевидно, что значительная часть инжектированных дырок не попадет на эмиттер и будет рекомбинировать на поверхности, прилегающей к эмиттеру, а также в объеме пассивной области базы.

Это означает уменьшение эффективного времени жизни и диффузионной длины:

$$\tau_I < \tau_N; \quad L_I < L_N.$$

Соответственно согласно (4-20) уменьшится величина α , т. е.

$$\alpha_I < \alpha_N.$$

Практически коэффициент α_I лежит в пределах 0,75—0,9. Поскольку в инверсном режиме траектории дырок, движущихся от коллектора к эмиттеру, в среднем длиннее, чем ω (из-за движения части дырок через промежуточную область), следует согласно (4-44) и (4-48) ожидать ухудшения переходных и частотных свойств:

$$\tau_{aI} > \tau_{aN}; \quad \omega_{aI} < \omega_{aN},$$

что имеет место в действительности.

4.6. ЗАВИСИМОСТЬ ПАРАМЕТРОВ ОТ РЕЖИМА И ТЕМПЕРАТУРЫ

Зависимость от режима. Параметры транзистора зависят от рабочей точки. Величинами, определяющими режим, считаются ток I_b и напряжение U_k . Посмотрим, как меняются параметры транзистора в зависимости от этих величин.

Коэффициент передачи тока α согласно (4-20) зависит от напряжения U_k из-за модуляции толщины базы. Чем больше (по модулю) коллекторное напряжение, тем уже база и тем ближе к единице коэффициент переноса дырок. Следовательно, коэффициент α увеличивается с ростом U_k , приближаясь к величине γ . Вторым фактором, приводящим к зависимости $\alpha(U_k)$, является ударная ионизация в коллекторном переходе. При наличии ионизации ток I_k возрастает в M раз, где M определяется выражением (2-55). Соответственно получаем:

$$\alpha_M = M\alpha,$$

где $\alpha = \gamma i$. Очевидно, что даже очень незначительное превышение M над единицей может существенно приблизить коэффициент α к единице. Более того, при сравнительно небольшом напряжении (20—30 В) может получиться $\alpha > 1$, что качественно меняет свойства транзистора (см. § 5-2).

Поскольку относительные изменения α очень невелики и их трудно отобразить на графике, на рис. 4-19, а показана кривая $1/(1-\alpha)$, которая гораздо ярче иллюстрирует эти изменения, сохраняя качественный характер зависимостей.

Зависимость α от тока эмиттера (рис. 4-19, б) обусловлена главным образом изменением коэффициента инжекции γ : с увеличением тока модулируется (уменьшается) сопротивление базы и согласно (2-74) уменьшается коэффициент γ . Соответствующий спад коэффициента α является важным фактором, ограничивающим максимальный рабочий ток транзистора. Уменьшение коэф-

фициента α в области малых токов объясняется двумя причинами. Во-первых, уменьшается эквивалентный коэффициент диффузии [см. (2-72)]. Во-вторых, возрастает роль тока рекомбинации в области эмиттерного перехода (см. § 2-8). Поскольку этот ток обусловлен уходом электронов из базы, общий коэффициент инжекции соответственно уменьшается.

Первая причина [63], свойственная всем бездрейфовым транзисторам, приводит к уменьшению α всего на несколько процентов. Вторая причина, характерная для кремниевых транзисторов [64], может привести к уменьшению α в 2 раза и больше (при этом величина $1/(1-\alpha)$ может уменьшиться в десятки раз).

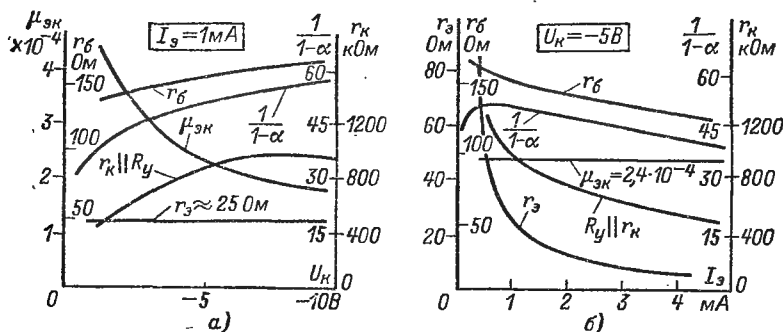


Рис. 4-19. Зависимость статических параметров транзистора от режима.

а — при постоянном токе эмиттера; б — при постоянном коллекторном напряжении.

В настоящее время, особенно в связи с разработкой маломощных интегральных схем, поведение транзисторов в области малых токов и напряжений привлекает особое внимание. Вопрос этот в общем достаточно сложный, но некоторое представление о зависимости $\alpha(I_E)$ при малых токах можно получить из следующих соображений.

Будем считать, что электронная составляющая эмиттерного тока обусловлена только рекомбинацией в переходе и, следовательно, имеет структуру (2-61):

$$I_{эн} = a_n e^{U/2\varphi T},$$

а дырочная составляющая имеет обычную структуру (2-33):

$$I_{эп} = a_p e^{U/\varphi T},$$

где a_p определяется выражением (2-366). Запишем коэффициент инжекции в виде

$$\gamma = \frac{I_{эп}}{I_{эп} + I_{эн}} = \frac{1}{1 + a e^{-U/2\varphi T}}.$$

Если экспоненциальный множитель выразить через $\sqrt{T_{эп}}$ и, полагая $\gamma \approx 1$, заменить $I_{эп}$ на I_s , то

$$\gamma \approx \frac{1}{1 + b \sqrt{T_s}},$$

где параметр b после подстановки и преобразований имеет следующее значение:

$$b = \frac{\varphi_T}{\tau} \sqrt{\frac{2\epsilon_0 \epsilon W S}{\Delta \varphi_0 D}}.$$

В области микротоков можно принять $\alpha = \gamma$, а величину $1/(1 - \alpha)$ (рис. 4-19), которая более наглядно характеризует изменения α , представить в виде

$$\frac{1}{1 - \alpha} = 1 + \frac{1}{b} \sqrt{I_3} \approx \sqrt{I_3} / b.$$

Приведенные соображения подтверждают и конкретизируют спад α с уменьшением тока, а также иллюстрируют зависимость этого спада от ряда факторов. В первую очередь следует отметить прямую зависимость α от времени жизни τ и тем самым от свойств и состояния поверхности в области перехода.

Спад коэффициента передачи в области больших и малых токов приводит к наличию максимума на кривой α , который имеет место при некотором небольшом токе. Этот ток обычно близок к рекомендуемому в качестве номинального.

Сопротивление эмиттерного перехода r_e , как следует из выражения (4-22), обратно пропорционально току эмиттера вплоть до очень малых значений последнего. Зависимость $r_e(U_k)$ очень слаба, и практически ею можно пренебречь.

Сопротивление коллекторного перехода r_k согласно (4-24) тоже обратно пропорционально току I_3 , но существенно зависит и от напряжения U_k , а именно пропорционально $\sqrt{U_k}$. С ростом U_k от 1 до 10—20 В сопротивление r_k должно было бы увеличиться в 3—4 раза. Однако этот эффект обычно маскируется поверхностными утечками, а также ударной ионизацией в коллекторном переходе. В результате сопротивление r_k оказывается меньше ожидаемого значения и с ростом U_k даже уменьшается.

Чтобы проиллюстрировать влияние ударной ионизации на величину r_k , положим $(U_k/U_M)^n \ll 1$ и запишем коэффициент ударной ионизации (2-55) в приближенном виде:

$$M \approx 1 + (U_k/U_M)^n.$$

Если теперь подставить $\alpha_M = M\alpha$ в выражение (4-8) и продифференцировать его по U_k , считая α постоянным, получим:

$$r_k = \frac{U_k}{nI_k} \left(\frac{U_M}{U_k} \right)^n. \quad (4-67)$$

Пусть, например, $U_M = 50$ В; $I_k = 1$ мА; $U_k = 5$ В; $n = 3$; при этом $r_k = 1,7$ МОм, тогда как из формулы (4-24) для того же режима $r_k = 4,2$ МОм.

Коэффициент обратной связи $\mu_{эк}$ согласно (4-25) обратно пропорционален $\sqrt{U_k}$ и не зависит от тока I_3 .

Объемное сопротивление базы r_6 модулируется при больших токах эмиттера. Эта модуляция имеет место в первую очередь в активной области базы. Поэтому с ростом тока I_3 сопротивление активной области играет все меньшую роль и суммарная

величина r_6 все больше определяется пассивными областями [см. рис. 4-14 и формулу (4-27)]. Зависимость сопротивления активной области базы r_{61} от эмиттерного тока примерно такая же, как в случае диода с тонкой базой (см. рис. 2-34, кривая $w/L \leq 0,5$). Если при малом токе составляющая r_{61} имеет значительный удельный вес в общем сопротивлении r_6 , то с ростом тока значение r_6 может измениться весьма существенно (в 2—3 раза и больше).

Изменение коллекторного напряжения приводит к модуляции толщины базы и обуславливает зависимость $r_6(U_K)$. Эта зависимость относится тоже в основном к активной области и выражается в увеличении составляющей r_{61} с ростом напряжения U_K . Обычно зависимость $r_6(U_K)$ выражена слабо, за исключением высокочастотных транзисторов с особо тонкой базой.

Время жизни τ с увеличением уровня инжекции (т. е. тока I_B) меняется по-разному в зависимости от удельного сопротивления базы (рис. 1—26, а). Влияние коллекторного напряжения на время жизни практически отсутствует.

Время диффузии t_D и пропорциональная ему постоянная времени τ_a зависят от тока в той мере, в какой коэффициент диффузии зависит от уровня инжекции. Практически в области больших токов наблюдается некоторое увеличение времени t_D [65]. Коллекторное напряжение влияет на параметры t_D и τ_a благодаря модуляции толщины базы. А именно с ростом U_K толщина w уменьшается, что уменьшает параметры t_D и τ_a иногда на десятки процентов (если база достаточно тонка). Граничная частота f_a соответственно возрастает.

Что касается барьерных емкостей, то их зависимость от режима определяется формулами (2-82)

Зависимость от температуры. Параметры транзистора зависят от температуры даже при неизменной рабочей точке ($I_B = \text{const}$; $U_K = \text{const}$).

Коэффициент передачи α согласно (4-20) и (2-35) зависит от температуры через параметры τ и ρ_6 (величину ρ_6 можно считать постоянной, так как слой эмиттера является полуметаллом). Из этих параметров главную роль играет время жизни, которое существенно возрастает с температурой (см. рис. 1-26, б). Поэтому коэффициент α растет при нагреве транзистора и уменьшается при его охлаждении. На рис. 4-20 для большей ясности показана температурная зависимость величины $1/(1 - \alpha)$.

Сопротивление эмиттерного перехода r_e согласно (4-22) линейно зависит от температуры через параметр Φ_T . Легко показать, что значение r_e меняется приблизительно на $0,33\%/^{\circ}\text{C}$.

Сопротивление коллекторного перехода r_K согласно (4-24) зависит от температуры в основном через диффузионную длину L (т. е. через время жизни) и должно увеличиваться при нагреве транзистора. Такое увеличение действительно наблюдается в области отрицательных температур, но в районе комнатной температуры (а иногда и раньше) оно сменяется

спадом и кривая $r_k(T)$ имеет максимум. Уменьшение r_k при повышении температур объясняется влиянием утечек, а также изменением коэффициента ударной ионизации [см. (4-67)1].

Коэффициент обратной связи $\mu_{\text{ЭК}}$ согласно (4-25) зависит от температуры через параметр Φ_T , т. е. линейно, как и сопротивление r_b .

Объемное сопротивление базы r_b меняется с температурой постольку, поскольку меняется удельное сопротивление. Зависимость $\rho(T)$ в общем случае, как известно, нелинейна (см. § 1-9); во многом она зависит от концентрации примесей в базовом слое. В случае относительно низкоомной базы ($\rho_b \ll \rho_i$), что характерно для кремниевых транзисторов, сопротивление r_b монотонно возрастает в рабочем диапазоне температур. В случае германиевых транзисторов база нередко бывает относительно высокоомной ($\rho_b \lesssim \rho_i$); тогда сопротивление r_b имеет максимум при температуре 20—70°С (рис. 4-20), после чего уменьшается, поскольку примесный полупроводник постепенно превращается в собственный.

Время жизни τ , как уже отмечалось, увеличивается с ростом температуры (см. рис. 1-26, б), особенно тогда, когда сопротивление базы приближается к собственному. Поэтому такая зависимость сильнее выражена у германиевых транзисторов.

Время диффузии t_D и постоянная времени τ_a зависят от температуры через коэффициент диффузии D и поэтому согласно (1-75а) должны несколько увеличиваться при нагреве транзистора. Поскольку эта зависимость имеет вид $\sqrt{T}$, она обычно не играет практической роли.

Для германиевых транзисторов важное значение имеет температурная зависимость теплового тока коллектора I_{k0} . Будучи очень небольшим при комнатной температуре (0,1—0,5% режимного тока I_a), он сильно возрастает при нагреве транзистора, а это, как видно из рис. 4-11, а, смещает все кривые коллекторного семейства характеристик. В результате получается косвенная температурная зависимость параметров. Функцию $I_{k0}(T)$ нет необходимости рассматривать подробно, так как она проанализирована в § 2-6 применительно к тепловому току диода. Напомним только, что эта функция имеет экспоненциальный характер и что ток I_{k0} у германиевых транзисторов удваивается с ростом температуры на 8—10°С.

У кремниевых транзисторов при температуре до 100°С главную роль играет не тепловой ток I_{k0} , а ток термогенерации I_G , температура удвоения которого составляет около 10°С. Однако с токами I_{k0} и I_G в кремниевых транзисторах, как правило, можно не считаться из-за их малости даже при весьма высокой температуре.

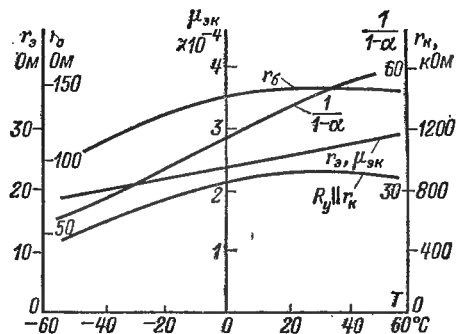


Рис. 4-20. Зависимость статических параметров транзистора от температуры.

Полная аналогия с диодами имеет место в отношении температурной зависимости напряжения на эмиттерном переходе при заданном токе I_B (см. § 2-8). Напомним, что зависимость эта почти линейна (при низких уровнях инжекции) и характеризуется отрицательной чувствительностью, близкой к $2 \text{ мВ/}^\circ\text{С}$ для обоих типов транзисторов — кремниевых и германиевых.

4-7. ХАРАКТЕРИСТИКИ И ПАРАМЕТРЫ ТРАНЗИСТОРА ПРИ ВКЛЮЧЕНИИ С ОБЩИМ ЭМИТТЕРОМ

Статические характеристики и параметры. При включении ОЭ (см. рис. 4-4, б) входным током является ток базы, который и прием за параметр коллекторного семейства характеристик.

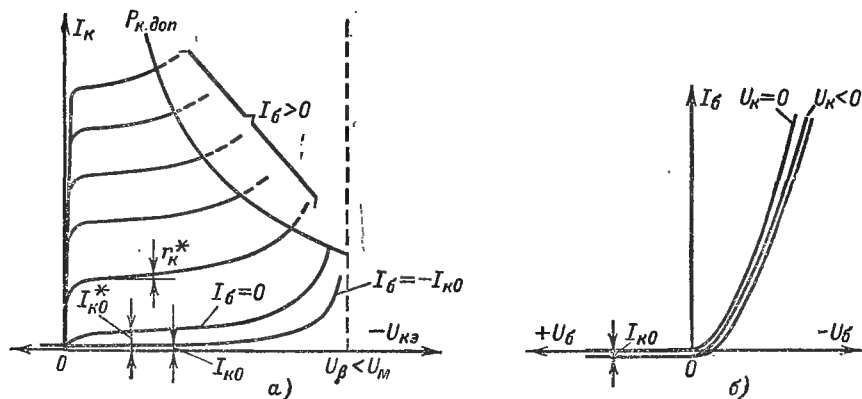


Рис. 4-21. Статические характеристики транзистора при включении по схеме с общим эмиттером.

а — выходные; б — входные.

Характеристики, показанные на рис. 4-21, несколько отличаются от характеристик ОБ. Главные отличительные черты включения ОЭ, вытекающие из сравнения рис. 4-21 и 4-11, сводятся к следующему.

а) Кривые коллекторного семейства не пересекают ось ординат и полностью расположены в 1-м квадранте. Это легко понять из соотношения $|U_{КЭ}| = |U_{КБ}| + U_B$: кривые ОЭ получаются путем сдвига кривых ОБ (см. рис. 4-11, а) на величину U_B , которая тем больше, чем больше ток.

б) Кривые коллекторного семейства менее регулярны, чем в схеме ОБ: они имеют гораздо больший, неодинаковый наклон и заметно сгущаются при больших токах. Ток при оборванной базе (когда $I_Б = 0$) намного больше тока I_{K0} при оборванном эмиттере и зависит от выходного напряжения. Входной ток $I_Б$ в принципе может иметь не только положительную, но и небольшую отрицательную величину (т. е. может втекать в базу).

в) Напряжение пробоя, которое на рис. 4-21, a обозначено через U_{β} , меньше, чем в схеме ОБ, по причинам, которые будут рассмотрены ниже.

Влияние коллекторного напряжения на ток базы (рис. 4-21, b) можно понять из того, что приращение $\Delta U_{кз}$, во-первых, вызывает приращение э. д. с. обратной связи $\mu_{эк} \Delta U_{к}$ и, во-вторых, частично падает на эмиттерном переходе. Оба эти влияния направлены навстречу друг другу, причем главную роль играет второе. Поэтому с ростом (по модулю) напряжения $U_{кз}$ кривые на рис. 4-21, b сдвигаются вправо — в сторону больших напряжений $U_{б}$.

В аналитическом виде семейство коллекторных характеристик ОЭ при работе транзистора в активном режиме ($U_{к} < 0$) получается из выражения (4-10), если в правую часть подставить очевидное соотношение $I_{э} = I_{к} + I_{б}$ и выразить ток коллектора через ток базы:

$$I_{к} = \frac{\alpha}{1-\alpha} I_{б} + \frac{I_{к0}}{1-\alpha} + \frac{U_{к}}{\bar{r}_{к}(1-\alpha)}.$$

Коэффициент при токе $I_{б}$ является *интегральным коэффициентом передачи базового тока*. Для этого важнейшего параметра введем специальное обозначение:

$$\beta = \frac{\alpha}{1-\alpha}. \quad (4-68)$$

Тогда функцию $I_{к}(I_{б})$ можно записать в следующей форме, аналогичной форме выражения (4-10):

$$I_{к} = \beta I_{б} + I_{к0}^* + \frac{U_{к}}{\bar{r}_{к}^*}, \quad (4-69)$$

где

$$I_{к0}^* = (1 + \beta) I_{к0}; \quad (4-70)$$

$$\bar{r}_{к}^* = \frac{\bar{r}_{к}}{1 + \beta}. \quad (4-71)$$

Часто последним членом в (4-69) пренебрегают; тогда получается аналог выражения (4-8) *:

$$I_{к} = \beta I_{б} + I_{к0}^*. \quad (4-72)$$

Минимальное значение коллекторного тока ($I_{к} = I_{к0}$) получается при токе базы $I_{б} = -I_{к0}$. Следовательно, в диапазоне от $I_{б} = 0$ до $I_{б} = -I_{к0}$ транзистор в схеме ОЭ может управляться отрицательным входным током, что уже отмечалось; однако такая возможность относится только к германиевым транзисторам, поскольку у кремниевых транзисторов ток $I_{к0}$ практически равен нулю.

Величина β , входящая в формулы (4-69) и (4-72), является интегральной, так как связывает между собой полные токи $I_{к}$ и $I_{б}$.

* Такое пренебрежение не всегда оправдано, поскольку при этом не учитывается наклон кривых на рис. 4-21, a . Током $I_{к0}^*$ можно пренебрегать только у кремниевых транзисторов и то при не очень высокой температуре (до 60—80 °C).

Из формулы (4-72) легко получить определение интегрального коэффициента передачи [ср. с (4-12а)]*:

$$\bar{\beta} = \frac{I_K - I_{K0}}{I_6 + I_{K0}}. \quad (4-73a)$$

Дифференциальный коэффициент передачи определяется по аналогии с (4-11а) как

$$\beta = \left. \frac{dI_K}{dI_6} \right|_{U_K = \text{const}}. \quad (4-73b)$$

Учитывая соотношение $dI_6 = dI_9 - dI_K$, убеждаемся, что формула (4-68) действительна не только для интегрального, но и для дифференциального коэффициента передачи. Связь между β и $\bar{\beta}$ устанавливается тем же путем, что и в случае α , и имеет вид [ср. с (4-12б)]:

$$\beta = \bar{\beta} + (I_6 + I_{K0}) \frac{d\bar{\beta}}{dI_6}. \quad (4-74)$$

В дальнейшем мы пренебрежем зависимостью β от тока, за исключением специальных случаев, и будем использовать обозначение β как для дифференциальной, так и для интегральной величин.

Учитывая, что $\alpha \approx 1$, из (4-68) легко сделать вывод, что $\beta \gg 1$. Например, при типичных значениях $\alpha = 0,98 \div 0,99$ получается $\beta = 50 \div 100$. Поскольку разность $1 - \alpha$ очень мала, ясно, что коэффициент β гораздо сильнее зависит от всех тех факторов, от которых зависит коэффициент α . К числу этих факторов относятся в первую очередь режим, температура и частота. Зависимость коэффициента усиления от частоты подробно рассматривается в следующем разделе. Что касается зависимостей от режима и температуры — то они делаются особенно наглядными, если в формулу (4-68) подставить α из (4-21а).

Тогда в общем случае

$$\beta = \frac{\gamma}{1 - \gamma + \frac{1}{2} (\omega/L)^2}. \quad (4-75a)$$

При условии $1 - \gamma \ll \frac{1}{2} (\omega/L)^2$ получается:

$$\beta \approx 2 \frac{L^2}{\omega^2} = \frac{2D\tau}{\omega^2} = \frac{\tau}{t_D}. \quad (4-75b)$$

Зависимость β от тока коллектора скрыта в параметрах γ и τ , зависимость от напряжения коллектора — в параметре ω , а зависимость от температуры — в параметрах D и τ . Все эти зависимости по существу определяются кривыми на рис. 4-19 и 4-20, где использована величина $1/(1 - \alpha) \approx \beta$.

* Для кремниевых транзисторов

$$\bar{\beta} = I_K / I_6,$$

поскольку током I_{K0} обычно можно пренебречь.

Нетрудно убедиться, что формула (4-72) соответствует эквивалентная схема для постоянных составляющих, показанная на рис. 4-22 и являющаяся аналогом эквивалентной схемы ОБ на рис. 4-12, б. При необходимости обе схемы можно дополнить соответственно сопротивлениями r_k и r_k^* . И та, и другая схемы не содержат генератора $\mu_{эк} U_k$ и потому не отражают реального сдвига входных характеристик в зависимости от U_k , что практически не существенно.

Теперь рассмотрим специфику характеристик ОЭ в области пробоя.

При наличии ударной ионизации коэффициент передачи α увеличивается в M раз. Подставляя значение $M\alpha$ в формулу (4-68), приходим к выводу, что при $M\alpha \rightarrow 1$ коллекторный ток в схеме ОЭ неограниченно увеличивается. Следовательно, напряжение лавинного пробоя в схеме ОЭ определяется условием

$$M\alpha = 1; \quad (4-76)$$

с учетом соотношения (2-55) оно оказывается равным:

$$U_B = U_M \sqrt{1 - \alpha}. \quad (4-77)$$

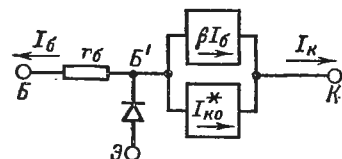


Рис. 4-22. Эквивалентная схема ОЭ для постоянных составляющих.

Индекс β принят здесь потому, что при этом напряжении результирующий коэффициент передачи базового тока обращается в бесконечность:

$$\beta = \frac{M\alpha}{1 - M\alpha} = \infty.$$

Для типичных значений $\alpha = 0,98$ и $n = 3$ получается $U \approx 0,3 U_M$. Как видим, пробой в схеме ОЭ происходит при значительно меньшем напряжении, чем в схеме ОБ. Это обстоятельство отражается в справочных данных [12, 66], в которых допустимое значение $(U_{кэ})_{доп}$ всегда в 2—3 раза меньше допустимого значения $(U_{кб})_{доп}$. Следует, однако, иметь в виду, что такая разница предполагает заданный ток базы, т. е. бесконечно большое сопротивление в цепи базы. При малом сопротивлении $R_б$ разница уменьшается, а при $R_б = 0$ она практически исчезает.

Второй тип пробоя в схеме ОЭ является еще более специфичным и носит название *эффекта смыкания*. В основе этого эффекта лежит расширение коллекторного перехода при увеличении напряжения U_k . Если слой базы достаточно тонкий, а напряжение U_k достаточно велико, то эмиттерный и коллекторный переходы могут «сомкнуться», т. е. толщина базы w делается равной нулю. При этом согласно (4-18) коэффициент переноса дырок делается равным единице. Соответственно резко возрастает коэффициент β , и практически имеет место пробой транзистора.

Если принять, что переход сосредоточен в базе, то напряжение смыкания U_w определяется из условия

$$l_k = w_0,$$

где l_k — ширина коллекторного перехода, а w_0 — толщина базы при нулевом коллекторном напряжении. Используя выражение (2-12) и переходя от концентрации примеси к удельному сопротивлению, нетрудно представить напряжение смыкания в следующем виде:

$$U_w = \frac{w_0^2}{2\epsilon_0 \epsilon_r \mu_6}. \quad (4-78)$$

Если, например, $\rho = 20 \text{ Ом} \cdot \text{см}$ и $w_0 = 20 \text{ мкм}$, то для электронного германия получим $U_w \approx 20 \text{ В}$. В зависимости от того, какая из величин (U_β или U_w) меньше, пробой будет иметь тот или иной характер. Внешнее проявление пробоя в обоих случаях одинаково, поэтому при $U_w < U_\beta$ достаточно заменить на рис. 4-21 критическое напряжение U_β на U_w .

Нетрудно заметить, что механизмы лавинного пробоя и смыкания тоже одинаковы и характеризуются выполнением условия (4-76). Анализируя лавинный пробой, мы считали $\alpha = \text{const}$, т. е. пренебрегали зависимостью $\alpha(U_k)$. Анализируя смыкание, мы считали $M = 1$, т. е. пренебрегали ударной ионизацией. Учитывая обе зависимости $\alpha(U_k)$ и $M(U_k)$, можно получить из условия (4-76) общее выражение для напряжения пробоя в схеме ОЭ. В частных случаях оно переходит в выражения (4-77) и (4-78).

Такой общий анализ затруднен, если показатель n в формуле (4-77) не равен 2, так как при этом получается неквадратное уравнение относительно напряжения. Но даже в случае квадратного уравнения напряжение пробоя выражается громоздкой формулой. Не приводя этой формулы, ограничимся критерием лавинного пробоя, который из нее вытекает:

$$w_0 \gg \frac{\epsilon_0 \epsilon_r \mu U_M}{L}. \quad (4-79)$$

При обратном (сильном) неравенстве имеет место смыкание. В промежуточном случае характер пробоя смешанный.

Поскольку величина U_M согласно (2-56) пропорциональна ρ , ясно, что лавинный пробой свойствен транзисторам со сравнительно низкоомным материалом базы. Пусть, например, в сплавном германиевом транзисторе $w_0 = 20 \text{ мкм}$; $L = 0,01 \text{ см}$ и $\rho = 2 \text{ Ом} \cdot \text{см}$; тогда согласно (2-56) $U_M = 120 \text{ В}$ и правая часть (4-79) составит около 1,5 мкм. При этом условие (4-79) выполняется и пробой является лавинным. Если же $\rho = 20 \text{ Ом} \cdot \text{см}$, то правая часть (4-79) равна 60 мкм и можно считать, что пробой обусловлен смыканием.

Динамические параметры. При включении транзистора по схеме ОЭ частотные и временные зависимости свойственны не только коэффициенту β , но и коллекторному сопротивлению, которое согласно (4-71) зависит от β .

Для выяснения инерционных свойств коэффициента β подставим в формулу (4-68) изображение $\alpha(s)$ из (4-43) или комплексную величину α из (4-46). Тогда после несложных преобразований полу-

$$\beta(s) = \frac{\beta}{1 + s\tau_\beta}; \quad (4-80a)$$

$$\dot{\beta} = \frac{\beta}{1 + j \frac{\omega}{\omega_\beta}}, \quad (4-80b)$$

где β определяется формулой (4-68), а постоянная времени τ_β и граничная частота ω_β — формулами

$$\tau_\beta = \frac{\tau_\alpha}{1 - \alpha} = (1 + \beta) \tau_\alpha; \quad (4-81a)$$

$$\omega_\beta = (1 - \alpha) \omega_\alpha = \frac{\omega_\alpha}{1 + \beta}. \quad (4-81b)$$

Если положить $\gamma = 1$ (т. е. $\alpha = \kappa$), то из выражений (4-35), (4-44) и (4-81a) получаем:

$$\tau_\beta = \tau. \quad (4-82)$$

Равенство (4-82) имеет достаточно общий характер и не связано с использованием 1-го приближения для $\alpha(s)$. Физические основы этого равенства следующие. Базовый ток всегда обусловлен основными носителями, в данном случае электронами. Задача базового тока — компенсировать убыль электронного заряда в базе, вызываемую двумя факторами: уходом электронов через эмиттерный переход¹ и рекомбинацией электронов в базе. При условии $\gamma = 1$ эмиттерный ток будет чисто дырочным и первый фактор отсутствует. Значит, при изменениях базового тока заряд в базе будет меняться только за счет рекомбинации, т. е. с постоянной времени τ . С такой же постоянной времени будет меняться и ток коллектора, поскольку он (а точнее, почти равный ему ток I_α) согласно (4-62) пропорционален заряду.

Условие $\gamma = 1$, лежащее в основе (4-82), нарушается при высоких уровнях инжекции [см. (2-74)], а также во время коротких переходных процессов [см. (4-306)]. В обоих случаях равенство (4-82) становится неточным.

Выражения (4-81) показывают, что переходные и частотные свойства в схеме ОЭ хуже, чем в схеме ОБ. В дальнейшем мы убедимся, что в реальных схемах этот недостаток в значительной мере сглаживается, тогда как преимущество схемы ОЭ — усиление входного тока — сохраняется.

Физические причины различия переходных и частотных свойств в схемах ОБ и ОЭ полезно проиллюстрировать с помощью векторных диаграмм (рис. 4-23). Эти диаграммы показывают, что даже небольшой сдвиг фаз между близкими по величине векторами $\dot{I}_\alpha$ и $\dot{I}_\kappa$ вызывает значительный сдвиг фаз между каждым из них и разностным вектором $\dot{I}_\beta$. Кроме того, величина последнего очень быстро возрастает с частотой в первую очередь из-за фазового сдвига между токами $\dot{I}_\alpha$ и $\dot{I}_\kappa$. Поэтому если даже модуль $\dot{I}_\kappa$ остается

¹ Коллекторный переход находится под отрицательным смещением, поэтому уход электронов через него исключен.

почти таким же, как и при низкой частоте (рис. 4-23, б), то модуль $\dot{I}_6$ уже сильно увеличивается, а следовательно, коэффициент β уменьшается [58].

Если в формулу (4-68) подставить не 1-е приближение (4-43), а аппроксимацию (4-49), то после ряда преобразований¹ изображение $\beta(s)$ можно записать в той же форме, что и (4-49):

$$\beta(s) = \frac{\beta e^{-st_{3\beta}}}{1 + s\tau_\beta}. \quad (4-83a)$$

Соответствующая частотная характеристика имеет вид:

$$\dot{\beta} = \frac{\beta e^{-j\omega t_{3\beta}}}{1 + j \frac{\omega}{\omega_\beta}}. \quad (4-83б)$$

Параметры в этих формулах имеют следующие значения:

$$\tau_\beta = \frac{\tau_\alpha + t_{3\alpha}}{1 - \alpha} = (1 + c)(1 + \beta)\tau_\alpha; \quad (4-84a)$$

$$\omega_\beta = \frac{1}{\tau_\beta} = \frac{\omega_\alpha}{(1 + c)(1 + \beta)}; \quad (4-84б)$$

$$t_{3\beta} = \frac{\tau_\alpha t_{3\alpha}}{\tau_\alpha + t_{3\alpha}} = \frac{c}{1 + c} \tau_\alpha. \quad (4-85)$$

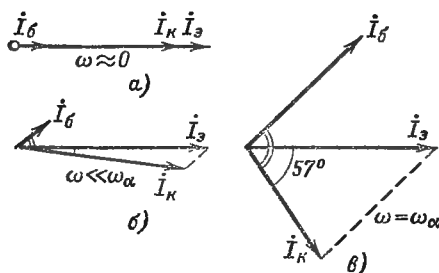
Легко заметить, что фактор задержки $t_{3\beta}/\tau_\beta$ несравненно (в десятки раз) меньше, чем $t_{3\alpha}/\tau_\alpha$. Следовательно, роль задержки в схеме ОЭ гораздо менее существенна, чем в схеме ОБ. Поэтому выражениями (4-83) редко пользуются на практике, за исключением тех случаев,

когда интересуются начальной стадией переходного процесса или очень высокими частотами, близкими к f_α . Наибольшее распространение имеют выражения (4-80), но параметры τ_β и ω_β обычно записываются в форме (4-84).

В области достаточно высоких частот ($\omega > 3\omega_\beta$) модуль коэффициента β согласно (4-80б) можно записать следующим образом:

$$|\dot{\beta}| \approx \frac{\beta}{\omega/\omega_\beta}.$$

Рис. 4-23. Векторные диаграммы схемы ОЭ при трех разных частотах.



Тем самым произведение $|\dot{\beta}|$ на частоту ω (в диапазоне $\omega > 3\omega_\beta$) оказывается постоянной величиной. Эту величину называют *предельной частотой коэффициента усиления тока* (ω_r), поскольку она

¹ Преобразования сводятся к следующему. Числитель и знаменатель умножаются на $(1 + st)e^{st_{3\alpha}}$, после чего экспонента разлагается в ряд с точностью до двух членов. Получающаяся функция вида (4-36а) приводится к форме (4-36б). Далее определяются соответствующие постоянные времени. Они оказываются резко различными, что дает основание рассматривать одну из них как постоянную времени τ_β , а другую — как задержку $t_{3\beta}$ (см. сноску на с. 199).

соответствует значению ω при $|\beta| = 1$:

$$\omega_T = \beta \omega_\beta = \frac{\alpha \omega_\alpha}{1 + c} = \frac{\alpha}{t_D}. \quad (4-86)$$

В последнее время в справочниках приводят обычно именно величину f_T , а не f_α (это связано с большим удобством измерения).

В тех случаях, когда существенна постоянная времени τ_V , нужно внести коррективы в формулы (4-84). Оценивая сумму постоянной времени τ_α и времени задержки $t_{3\alpha}$ с помощью выражений (4-57) и (4-58), приходим к выводу, что в любом случае (т. е. и при малых, и при больших значениях τ_V) постоянная времени τ_β имеет вид:

$$\tau_\beta = \frac{\tau_\alpha + t_{3\alpha} + \tau_V}{1 - \alpha} = \frac{t_D + r_3 C_3}{1 - \alpha}. \quad (4-87)$$

При этом для предельной частоты ω_T нетрудно получить обобщение формулы (4-86):

$$\frac{1}{\omega_T} = \frac{t_D + \tau_V}{\alpha} = \frac{1}{\omega_{i0}} + \frac{1}{\alpha} r_3 C_3. \quad (4-88)$$

Если снять экспериментальную зависимость $\omega_T(I_3)$ и построить ее в виде кривой $\omega_T(1/I_3)$, то экстраполяция к оси ординат (т. е. к значению $1/I_3 = 0$) дает величину ω_T , а наклон кривой к каждой точке определяет емкость C_3 [60].

Теперь рассмотрим коллекторное сопротивление схемы ОЭ, которое отличается от значения r_k . В самом деле, в схеме ОЭ (рис. 4-4, б) при заданном токе I_6 приращение выходного напряжения dU_{k3} распределяется между обоими переходами. В результате изменение тока I_k сопровождается таким же изменением тока I_3 и соответствующим изменением тока эквивалентного генератора αI_3 . Если положить $C_k = 0$, то согласно (4-10) приращение dI_k выразится операторной суммой

$$dI_k(s) = \frac{dU_k(s)}{r_k} + \alpha(s) dI_3(s),$$

где $dI_3 = dI_k$.

Отсюда изображение эквивалентного сопротивления коллекторного перехода в схеме ОЭ будет иметь вид:

$$r_k^*(s) = \frac{dU_k(s)}{dI_k(s)} = r_k [1 - \alpha(s)] = \frac{r_k}{1 + \beta(s)}. \quad (4-89)$$

При медленных изменениях коллекторного напряжения (или низкой частоте) вместо изображений можно использовать действительные величины; тогда сопротивление r_k^* выражается формулой, аналогичной (4-71):

$$r_k^* = r_k (1 - \alpha) = \frac{r_k}{1 + \beta}. \quad (4-90)$$

Сопротивление r_k^* в десятки раз меньше, чем сопротивление r_k (это уже отмечалось в связи с семейством характеристик на рис. 4-21, а).

Аналогичную методику можно использовать для определения коллекторной емкости в схеме ОЭ. Положим для простоты $r_k = \infty$. Тогда для переходных процессов роль сопротивления r_k будет играть емкостное сопротивление $X_c = 1/sC_k$. Заменяя в формуле (4-89) r_k на X_c , получаем изображение эквивалентной емкости:

$$C_k^*(s) = \frac{C_k}{1 - \alpha(s)} = C_k[1 + \beta(s)]. \quad (4-91)$$

При медленных изменениях тока (или низкой частоте) получаем действительную величину:

$$C_k^* = \frac{C_k}{1 - \alpha} = C_k(1 + \beta). \quad (4-92)$$

Емкость C_k^* в десятки раз больше емкости C_k .

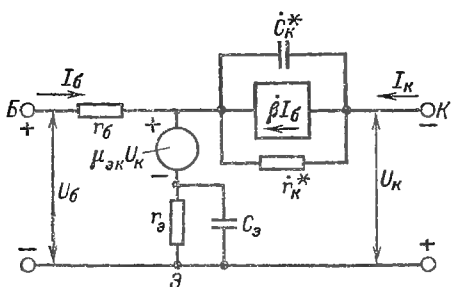
Заметим, что постоянная времени коллекторного перехода одинакова в схемах ОБ и ОЭ:

$$\tau_k = C_k r_k = C_k^*(s) r_k^*(s) = C_k^* r_k^*. \quad (4-93)$$

Например, при $r_k = 1$ МОм и $C_k = 5$ пФ получается $\tau_k = 5$ мкс.

Тот факт, что величины r_k^* и C_k^* , вообще говоря, являются спе-
раторными, имеет большое значение при анализе быстрых переход-
ных процессов и не может не учитываться.

Приведем еще выражение для операторного импеданса, состоя-
щего из параллельно соеди-
ненных сопротивления r_k^* и
емкости C_k^* :



$$Z_k^*(s) = r_k^* \frac{1 + s\tau_\beta}{(1 + s\tau_\alpha)(1 + s\tau_k)}. \quad (4-94a)$$

Здесь роль постоянной
времени τ_α обычно невелика,
и можно пользоваться упро-
щенным выражением

$$Z_k^*(s) \approx r_k^* \frac{1 + s\tau_\beta}{1 + s\tau_k}. \quad (4-94b)$$

Рис. 4-24. Эквивалентная схема ОЭ для переменных составляющих.

Эквивалентная схема для переменных составляющих после проведенного анализа и с учетом общей схемы (см. рис. 4-22) может быть изображена так, как показано на рис. 4-24.

Заметим, что в схеме ОЭ диффузионная емкость эмиттера определяется выражением (4-63а), а не (4-63б), т. е. в 1,5 раза больше, чем в схеме ОБ¹.

¹ Выражение (4-63а), как уже отмечалось, соответствует чисто экспоненциальному процессу изменения напряжения $U_э$. Именно такой процесс свойствен схеме ОЭ, в которой задержка фронта по отношению к длительности фронта имеет ничтожную величину [см. замечания к формуле (4-85)].

Схема с общим коллектором. Эта схема (см. рис. 4-4, в) имеет много общего со схемой ОЭ, потому что в обеих управляющим током является ток базы, а выходные токи ($I_{\text{в}}$ или $I_{\text{к}}$), как известно, различаются незначительно. Поэтому семейство выходных характеристик будет практически таким же, как на рис. 4-21, а, если ток коллектора заменить током эмиттера. Переходные и частотные свойства схемы ОК почти совпадают со свойствами схемы ОЭ, так как они определяются коэффициентом передачи

$$\frac{dI_{\text{в}}}{dI_{\text{б}}} = 1 + \beta,$$

мало отличающимся от β .

4-8. РАЗНОВИДНОСТИ ЭКВИВАЛЕНТНЫХ СХЕМ

Т-образные эквивалентные схемы, которые мы рассмотрели и которыми будем пользоваться в дальнейшем, не являются единственно возможными. В литературе можно встретить ряд других схем [58, 60], из которых чаще всего встречается П-образная эквивалентная схема. Кроме того, нередко (особенно в справочной литературе) транзистор рассматривается как бесструктурный четырехполосник с той или иной системой параметров.

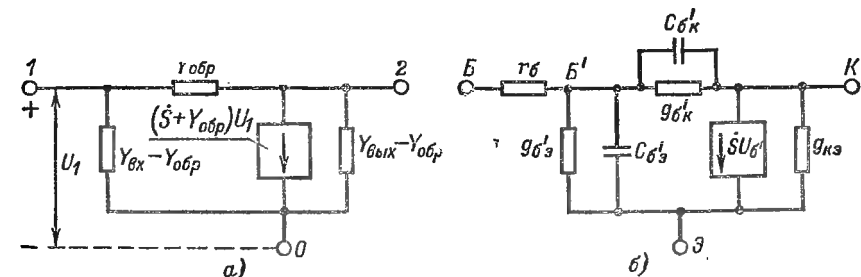


Рис. 4-25. П-образные эквивалентные схемы ОЭ.

а — основная; б — модифицированная (гибридная).

П-образные эквивалентные схемы. Основная П-образная схема для включения ОЭ показана на рис. 4-25, а а ее вариант, в котором выделено сопротивление базы, — на рис. 4-25, б. Как видим, в обеих схемах используются проводимости — комплексные (Y) или активные (g), а в качестве усилительного параметра используется крутизна S . Остановимся подробнее на «гибридной» П-образной схеме (рис. 4-25, б) как более распространенной и более специфичной для транзисторов.

Установим связь между параметрами П- и Т-образной схем (рис. 4-25, б и 4-24). Для этого сначала рассмотрим область низких частот, пренебрегая емкостями и частотной зависимостью параметров. Кроме того, исключим из рассмотрения сопротивление $r_б$, одинаковое для обеих схем. Тогда в каждой схеме останется по четыре

параметра: $g_{\delta'_{\delta}}, g_{\delta'_{\kappa}}, g_{\kappa\delta}, S$ и $r_{\delta}, \mu_{\delta\kappa}, r_{\kappa}^*, \beta$. Чтобы выразить одни параметры через другие, необходимы четыре уравнения. Такие уравнения получаются, если, например, приравнять входные (базовые) и выходные (коллекторные) токи в обеих схемах при заданном входном напряжении (между точками B' и δ) и коротком замыкании на выходе, а затем приравнять выходные (коллекторные) токи и входные (базовые) напряжения при заданном выходном напряжении и холостом ходе на входе. Точные решения несколько громоздки, поэтому имеет смысл принять некоторые несущественные ограничения, а именно: $r_{\delta} \ll r_{\kappa}^*, \mu_{\delta\kappa} \ll 1$. Тогда с учетом соотношения (4-26) связь между параметрами оказывается весьма простой:

$$S = \frac{\alpha}{r_{\delta}}; \quad (4-95a)$$

$$g_{\delta'_{\delta}} = \frac{1}{(1 + \beta) r_{\delta}}; \quad (4-95b)$$

$$g_{\delta'_{\kappa}} = \frac{\mu_{\delta\kappa} (1 - \alpha)}{r_{\delta}} = \frac{1}{2r_{\kappa}}; \quad (4-95b)$$

$$g_{\kappa\delta} = \frac{\mu_{\delta\kappa}}{r_{\delta}} = \frac{1}{2r_{\kappa}^*}. \quad (4-95r)$$

Подобно тому, как внутренние параметры Т-образной схемы связаны соотношением (4-26), параметры П-образной схемы, в чем легко убедиться, связаны соотношением

$$g_{\kappa\delta} g_{\delta'_{\delta}} = S g_{\delta'_{\kappa}}. \quad (4-96)$$

Теперь рассмотрим область высших частот, т. е. учтем на рис. 4-24 емкости C_{κ}^* и $C_{\delta, \delta}$ (последняя занимает то же место, что и C_{δ}), а на рис. 4-25, б — емкости $C_{\delta'_{\kappa}}$ и $C_{\delta'_{\delta}}$; кроме того, коэффициент β и крутизну S будем считать комплексными величинами. Тогда методика, аналогичная предыдущей, приводит к следующим результатам:

$$\dot{S} = \frac{\dot{\alpha}}{r_{\delta}} = \frac{\alpha/r_{\delta}}{1 + j\omega\tau_{\alpha}}; \quad (4-97a)$$

$$Y_{\delta'_{\delta}} = g_{\delta'_{\delta}} \frac{(1 + j\omega\tau_D)(1 + j\omega\tau_{\beta})}{1 + j\omega\tau_{\alpha}} \approx g_{\delta'_{\delta}} (1 + j\omega\tau_{\beta}); \quad (4-97b)$$

$$Y_{\delta'_{\kappa}} = \frac{1}{2Z_{\kappa}} = g_{\delta'_{\kappa}} (1 + j\omega\tau_{\kappa}); \quad (4-97b)$$

$$Y_{\kappa\delta} = \frac{1}{2Z_{\kappa}^*} \approx g_{\kappa\delta} \frac{1 + j\omega\tau_{\kappa}}{1 + j\omega\tau_{\beta}}. \quad (4-97r)$$

Структура проводимости $Y_{\delta'_{\kappa}}$ соответствует параллельному соединению низкочастотного сопротивления $2r_{\kappa}$ и емкости $C_{\kappa}/2$. Следовательно,

$$C_{\delta'_{\kappa}} = C_{\kappa}/2. \quad (4-98)$$

Структура проводимости $Y_{\delta'_{\delta}}$ соответствует параллельному соединению низкочастотного сопротивления $r_{\delta} (1 + \beta)$ и эквивалент-

ной емкости:

$$C_{6'3} = \frac{\tau_{\beta}}{r_s(1+\beta)} = \frac{t_D}{r_s} = C_{3.д} \quad (4-99)$$

Таким образом, в П-образной схеме в отличие от Т-образной частотная зависимость «сосредоточена» во в х о д н о й цепи ($\tau_{6'3} = \tau_{\beta}$), а параметр выходной цепи — крутизна S — зависит от частоты сравнительно слабо ($\tau_s = \tau_{\alpha}$).

Параметры основной П-образной схемы (рис. 4-25, а) в принципе нетрудно получить с помощью формул (4-97), учтя наличие сопротивления r_6 на входе. Однако параметры этой схемы оказываются частотно-зависимыми, что в общем случае весьма неудобно. Поэтому собственно П-образная схема используется лишь при анализе избирательных схем с практически постоянной рабочей частотой.

Параметры транзистора как четырехполюсника. Поскольку в Т- и П-образных схемах внутренняя базовая точка B' недоступна для подключения измерительных приборов, в справочной литературе обычно приводят параметры транзисторов, измеренные со стороны внешних зажимов. При этом транзистор рассматривается как «черный ящик», т. е. четырехполюсник с произвольной внутренней структурой. Четырехполюсник может характеризоваться различными системами параметров в зависимости от того, какие величины (из двух токов и двух напряжений) принимаются за аргументы, а какие — за функции.

Как известно, все системы параметров четырехполюсника в принципе равноценны, но по ряду причин большим распространением в транзисторной технике пользуется так называемая *смешанная h-система*, в которой на входе осуществляется режим холостого хода (задается режимный ток), а на выходе — режим короткого замыкания (задается режимное напряжение). Эта система соответствует принятым нами режимным величинам I_3 и $U_{к6}$ или I_6 и $U_{кз}$.

Система уравнений четырехполюсника в h -системе записывается следующим образом (рис. 4-26):

$$U_1 = h_{11}I_1 + h_{12}U_2; \quad (4-100a)$$

$$I_2 = h_{21}I_1 + h_{22}U_2. \quad (4-100б)$$

Отсюда видно, что параметр h_{11} есть входное сопротивление при коротком замыкании на выходе; h_{12} — коэффициент обратной передачи напряжения при холостом ходе на входе; h_{21} — коэффициент прямой передачи тока при коротком замыкании на выходе; h_{22} — выходная проводимость при холостом токе на входе. Эти определения лежат в основе соответствующих измерений.

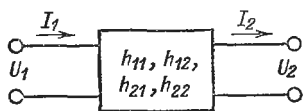


Рис. 4-26. Блок-схема четырехполюсника с h -параметрами.

Очевидно, что значения h -параметров в схемах ОБ и ОЭ будут различны. Установим связь параметров h_{11} , ..., h_{22} в схеме ОБ с параметрами Т-образной эквивалентной схемы на низких частотах, т. е. пренебрегая емкостями и комплексностью α на рис. 4-13. Задавая ток I_3 , полагая $U_K = 0$ на схеме рис. 4-13 и определяя величины U_3 и I_K , получаем:

$$\left| \frac{U_3}{I_3} \right|_{U_K=0} = h_{11} \approx r_3 + r_6(1 - \alpha); \quad (4-101a)$$

$$\left| \frac{I_K}{I_3} \right|_{U_K=0} = h_{21} = \alpha \frac{r_K}{r_K + r_6} \approx \alpha. \quad (4-101б)$$

Полагая теперь на рис. 4-13 ток $I_3 = 0$ (обрыв эмиттера по переменному току) и задавая U_K , получаем:

$$\left| \frac{U_3}{U_K} \right|_{I_3=0} = h_{12} \approx \frac{r_6}{r_K} + \mu_{ЭК}; \quad (4-101в)$$

$$\left| \frac{I_K}{U_K} \right|_{I_3=0} = h_{22} \approx \frac{1}{r_K}. \quad (4-101г)$$

С помощью формул (4-101) и учитывая соотношение (4-26), легко осуществить обратный расчет параметров α_K , r_K , r_3 , $\mu_{ЭК}$, r_6 по измеренным величинам h_{11} , ..., h_{22} :

$$\alpha \approx h_{21}; \quad (4-102a)$$

$$r_K \approx \frac{1}{h_{22}}; \quad (4-102б)$$

$$r_3 = 2 \left[h_{11} - \frac{h_{12}}{h_{22}} (1 - h_{21}) \right]; \quad (4-102в)$$

$$\mu_{ЭК} = \frac{h_{11}h_{22}}{1 - h_{21}} - h_{12}; \quad (4-102г)$$

$$r_6 = 2 \frac{h_{12}}{h_{22}} - \frac{h_{11}}{1 - h_{21}}. \quad (4-102д)$$

Частота переменных составляющих, при которой производятся измерения h -параметров этим методом, должна быть достаточно низкой (обычно до 500 Гц). Следует также заметить, что расчеты по формулам (4-102в) — (4-102д) оказываются весьма неточными, так как величины r_3 , $\mu_{ЭК}$ и r_6 определяются разностями почти одинаковых величин и малейшая погрешность при измерении h -параметров приводит к существенной погрешности в расчетах вплоть до получения отрицательных значений r_3 и r_6 . Поэтому сопротивление r_3 обычно рассчитывают по формуле (4-22), а сопротивление r_6 измеряют отдельно на достаточно высокой частоте (когда $1/\omega C_K \ll r_K$), подставляя $1/\omega C_K$ вместо r_K в формулу (4-101в) и считая $\mu_{ЭК} \ll \omega r_6 C_K$. Коллекторную емкость тоже измеряют отдельно с помощью куметра. Параметр h_{12} , измеренный на той же высокой частоте, позволяет непосредственно определить из (4-101в) постоянную времени $\tau_6 = r_6 C_K$, которая вместе с граничной частотой f_α характеризует частотные свойства транзистора [см. (4-65)]. Коэффициент $\mu_{ЭК}$ можно вычислить по формуле (4-26), зная величины r_3 , r_K и α .

На рис. 4-27 показана эквивалентная схема, соответствующая h -системе. Она имеет много общего с Т-образной схемой на рис. 4-13 и совпадает с ней (за исключением обозначений) при $r_6 = 0$, т. е. для идеального одномерного транзистора. Схема на рис. 4-27 сохра-

няет свою структуру при двух других включениях транзистора — ОЭ и ОК, хотя значения параметров при этом, конечно, меняются. Так, используя изложенную выше методику применительно к рис. 4-24, получаем для схемы ОЭ:

$$h_{11} \approx r_6 + (1 + \beta) r_3; \quad (4-103a)$$

$$h_{21} \approx \beta; \quad (4-103б)$$

$$h_{12} \approx \frac{r_3}{r_6^*} - \mu_{эк} = \frac{r_3}{2r_6^*}; \quad (4-103в)$$

$$h_{22} \approx \frac{1}{r_6^*}. \quad (4-103г)$$

Отсюда, как и в случае схемы ОБ, легко получить параметры Т-образной схемы по измеренным значениям h .

Сравнительная оценка. Из четырех рассмотренных эквивалентных схем (см. рис. 4-24, 4-25 и 4-27) имеет смысл сравнить только две: Т-образную (рис. 4-24) и гибридную П-образную (рис. 4-25, б). Остальные две неспецифичны для транзисторов и пригодны для любого активного четырехполюсника.

В литературе долгое время велись дискуссии по поводу относительных преимуществ Т- и П-образных схем. Сторонники П-образных схем (или как говорят, системы y -параметров) делали попытки приписать Т-образной схеме принципиально большую погрешность в области высоких частот [67, 68]. Последнее, разумеется, неверно, так как любые эквивалентные схемы с одинаковым числом степеней свободы (т. е. независимых параметров) равноценны, так же как четырехполюсники, описываемые разными системами параметров. Поэтому использование той или иной корректной схемы есть вопрос удобства или других приводящих обстоятельств, но не принципа.

Использование П-образных схем обычно мотивировалось желанием сблизить методы расчета транзисторных и ламповых схем [69]. При этом имели в виду, что основной параметр лампы — крутизна — имеет размерность проводимости, и поэтому стремились принять такую эквивалентную схему для транзистора, которая тоже состояла бы из проводимостей, а активные (усилительные) свойства которой характеризовались бы крутизной. Однако, помимо того, что лампы уже давно утратили свое былое значение в качестве универсальных элементов электронных схем, следует вспомнить, что анодная цепь лампы характеризуется не проводимостью, а внутренним сопротивлением. Кроме того, в ламповой технике, как известно, имеют широкое распространение понятия входного и выходного сопротивлений каскада, сопротивления нагрузки и др. Таким образом, говоря о «ламповой системе параметров», по существу имеют в виду только крутизну, а этого, конечно, недостаточно, чтобы специально приспособлять параметры транзистора к ламповым.

Транзистор, будучи прибором, качественно отличным от лампы, требует соответственно и особых количественных характеристик. К ним как раз и относятся Т-образные эквивалентные схемы, построенные на основе физических процессов в приборе, а не по какой-либо аналогии, которая по тем или иным причинам считается желательной, но может привести (и действительно приводила) к формальным результатам. В самом деле, если схема на рис. 4-25, б

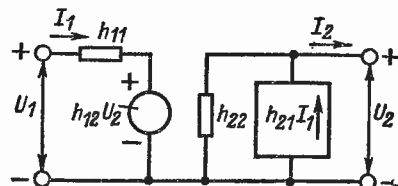


Рис. 4-27. Эквивалентная схема транзистора как четырехполюсника с h -параметрами.

еще сохраняя некоторую специфику транзистора, то «классическая» П-схема (см. рис. 4-25, а), к которой в конце концов переходят некоторые авторы [68, 69], полностью лишена этой специфики и применение ее для всех типов усилителей нельзя оправдать тем, что в некоторых частных случаях (нейтрализация резонансных усилителей) она оказывается несколько удобнее других. Что касается импульсных ключевых устройств, то для их анализа П-образная схема никем не используется; применять же для одного класса устройств (усилительных) одну эквивалентную схему, а для другого класса устройств (ключевых) — другую по меньшей мере неудобно. Заметим еще, что хотя параметры П- и Т-образной схем взаимосвязаны, но режимные и температурные зависимости параметров в последней проще, чем в первой. Достаточно указать на сложную структуру проводимости $g_{c'3}$, а также на режимную зависимость крутизны и диффузионной емкости [через сопротивление r_3 , см. (4-22)].

Приведенные соображения делают Т-образную схему предпочтительнее, хотя никаких принципиальных дефектов у П-образной схемы, разумеется, нет. Никакой принципиальной роли не играет также то обстоятельство, что у одной из схем управляющей величиной служит ток, а у другой — напряжение, хотя с измерительной точки зрения задать напряжения на $p-n$ переходе труднее, чем ток [см. замечания к формулам (2-37) и (2-38)], и это еще один аргумент в пользу Т-образной схемы.

4-9. СОБСТВЕННЫЕ ШУМЫ ТРАНЗИСТОРА

Собственные шумы в транзисторах, как и в других электронных приборах, разделяются на высокочастотные и низкочастотные [70]. Первые обусловлены дробовым эффектом и тепловыми флуктуациями в сопротивлениях слоев, вторые — поверхностными явлениями в области переходов.

Источники шумов. Д р о б о в о й ш у м, как известно, измеряется величиной

$$\overline{i_d^2} = 2qI \Delta f, \quad (4-104)$$

где I — ток, дискретная структура которого является причиной шума; Δf — полоса частот, в которой рассчитывается шум.

Из формулы (4-104) видно, что дробовой шум зависит только от д и а п а з о н а частот и не зависит от с р е д н е й частоты диапазона. Иначе говоря, спектр дробовых шумов является равномерным (так называемый «белый шум»).

Т е п л о в о й ш у м измеряется величиной

$$\overline{e_r^2} = 4kTR \Delta f, \quad (4-105)$$

где R — «шумящее» сопротивление; e_r — эквивалентная э. д. с., включенная последовательно с этим сопротивлением.

Из формулы (4-105) видно, что тепловой шум, как и дробовой, является «белым шумом», т. е. обладает равномерным спектром.

Н и з к о ч а с т о т н ы е (избыточные) шумы обычно обусловлены поверхностными явлениями и характерны тем, что их спектр неравномерен. В случае $\Delta f \ll f$ эти шумы можно записать в следующем виде:

$$\overline{i_n^2} = A \frac{\Delta f}{f^n}, \quad (4-106)$$

где A — коэффициент, величина и структура которого будут уточнены ниже. Показатель степени m в формуле (4-106) лежит в пределах 0,9—1,5, но, как правило, считают $m = 1$, что сильно упрощает расчеты.

Избыточный шум в области коллекторного перехода обусловлен флуктуациями поверхностной утечки. Поэтому он зависит от ширины перехода и тем самым от коллекторного напряжения. В первом приближении можно считать этот шум пропорциональным $\sqrt{U_K}$ [поскольку ширина ступенчатого перехода $i_K \sim \sqrt{U_K}$, см. (2-12)]:

$$\bar{i}_{п.к}^2 = A_K (U_K) \frac{\Delta f}{f} \approx A'_K \sqrt{U_K} \frac{\Delta f}{f}. \quad (4-107a)$$

Избыточный шум в области эмиттерного перехода обусловлен флуктуациями поверхностной рекомбинации инжектированных носителей. Поэтому он зависит от уровня инжекции, т. е. от эмиттерного тока. В первом приближении можно считать этот шум пропорциональным току эмиттера:

$$\bar{i}_{п.э}^2 = A_\varepsilon (I_\varepsilon) \frac{\Delta f}{f} \approx A'_\varepsilon I_\varepsilon \frac{\Delta f}{f}. \quad (4-107b)$$

У современных транзисторов избыточные шумы имеют существенное значение только при частотах менее нескольких килогерц. На частотах 10 кГц и выше с ними обычно можно не считаться.

Теперь покажем, как учитываются шумы транзистора на его эквивалентной схеме [71]. Возьмем за основу упрощенную схему на рис. 4-12, б исключим из нее сопротивление r_b (поскольку ему свойствен тепловой, а не дробовой шум) и разделим каждый из трех токов на составляющие (рис. 4-28). Смысл такого разделения в том, что каждая из составляющих имеет разное физическое происхождение (инжекция, экстракция, рекомбинация), а в этом случае, как показывает теория, нужно вычислять дробовые шумы раздельно. С электротехнической точки зрения указанные составляющие можно обеспечить источниками тока, подключенными между электродами Э, Б и К, как показано пунктиром на том же рисунке. Согласно (4-104) каждому такому источнику тока должен соответствовать генератор дробовых шумов, подключенный параллельно источнику тока. Кроме того, между электродами Э и Б' следует подключить генератор поверхностных шумов $\bar{i}_{п.э}$.

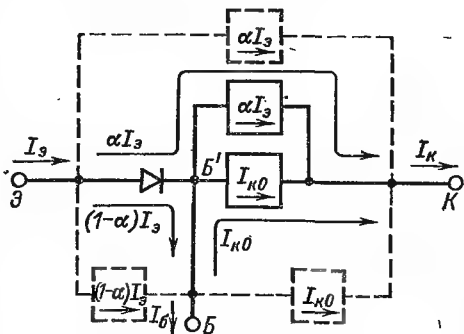


Рис. 4-28. Эквивалентная схема идеализированного транзистора. Пунктиром показаны токи — источники дробовых шумов.

а между электродами B' и K — генератор поверхностных шумов $\bar{i}_{п.к.}$. Тепловой шум в сопротивлении r_6 учитывается с помощью генератора шумов $\bar{e}_{т.6}$, включенного последовательно с r_6 . В результате малосигнальная эквивалентная схема транзистора с источниками шумов будет такой, как показано на рис. 4-29*. Здесь помимо собственных шумов учтен также тепловой шум в сопротивлении генератора (источника) сигнала R_r . Параметры шумовых генераторов на рис. 4-29 имеют следующие значения:

$$\bar{e}_{т.г}^2 = 4kTR_g \Delta f; \quad (4-108a)$$

$$\bar{e}_{т.6}^2 = 4kTr_6 \Delta f; \quad (4-108б)$$

$$\bar{i}_{эб}^2 = \left[2q(1 - \alpha) I_э + \frac{A_э}{f} \right] \Delta f; \quad (4-108в)$$

$$\bar{i}_{6'к}^2 = \left[2qI_{к0} + \frac{A_к}{f} \right] \Delta f; \quad (4-108г)$$

$$\bar{i}_{эк}^2 = 2q(\alpha I_э) \Delta f. \quad (4-108д)$$

Коэффициент шума. Уровень шумов в любом четырехполюснике, в том числе в транзисторном каскаде, можно оценивать величиной коэффициента шума¹.

Коэффициентом шума называется отношение полной мощности шумов в нагрузке к той части полной мощности, которая обусловлена тепловыми флуктуациями в сопротивлении генератора сигнала:

$$F_{ш} = \frac{(P_{ш})_{полн}}{(P_{т})_г}. \quad (4-109)$$

Из такого определения следует, что для идеального «нешумящего» четырехполюсника (или каскада) коэффициент шума равен единице, поскольку в этом случае шумы обусловлены только сопротивлением источника сигнала, т. е.

$(P_{ш})_{полн} = (P_{т})_г$.

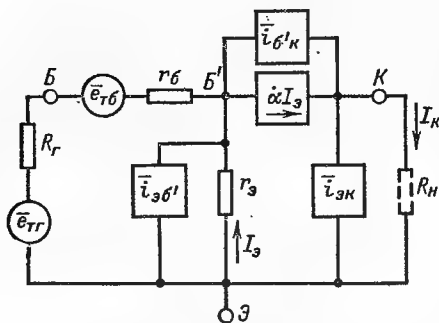


Рис. 4-29. Эквивалентная схема транзистора с генераторами шумов.

* На этой схеме, действительной для переменных составляющих, естественно, опущен генератор постоянного тока $I_{к0}$, а эмиттерный диод заменен дифференциальным сопротивлением $r_э$. Кроме того, принято включение ОЭ, наиболее распространенное на практике, причем в отличие от схемы на рис. 4-24 используется эквивалентный генератор $\alpha I_э$, с тем чтобы сохранить связь с рис. 4-28.

¹ Этот сравнительный параметр оказывается удобным и наиболее распространен в радиотехнике. В физике полупроводниковых приборов за последнее время считается более информативной абсолютная характеристика шумов с помощью двух эквивалентных шумовых генераторов (э. д. с. и тока), включенных на входе четырехполюсника.

Выведем коэффициент шума для простейшей схемы ОЭ, показанной на рис. 4-29. Сопротивление нагрузки R_H для простоты примем идеальным, т. е. пренебрежем его собственными шумами. Тогда все слагаемые полной мощности выходных шумов можно будет представить в форме $\bar{i}_m^2 R_H$. После подстановки в выражение (4-109) множитель R_H сократится и коэффициент шума выразится следующим образом:

$$F_{ш} = 1 + \frac{\sum_{m=2}^M \bar{i}_m^2}{\bar{i}_1^2}. \quad (4-110)$$

Здесь $\bar{i}_1$ — компонент тока I_K , пропорциональный э. д. с. $\bar{e}_{т.г.}$, а $\bar{i}_m$ — компоненты тока I_K , пропорциональные э. д. с. $e_{т.б}$ и токам других генераторов, показанных на рис. 4-29.

Чтобы определить компоненты $\bar{i}_m$, необходимо учесть действие эквивалентного генератора αI_3 . Покажем это на примере компонента $\bar{i}_1$. Соответствующая эквивалентная схема приведена на рис. 4-30, а, где α — модуль коэффициента передачи¹.

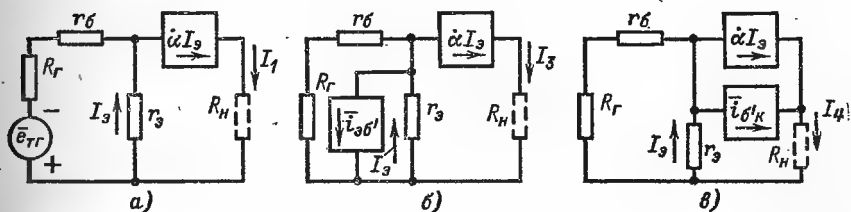


Рис. 4-30. Эквивалентные схемы для расчета компонентов полного тока шумов.

Воспользуемся *методом суперпозиции*. Оборвем сначала цепь генератора αI_3 . Тогда ток эмиттера будет равен:

$$I'_3 = \frac{\bar{e}_{т.г.}}{R_Г + r_б + r_з}.$$

Теперь, наоборот, закоротим источник э. д. с. $\bar{e}_{т.г.}$; тогда ток эмиттера опре-

$$I''_3 = \alpha I_3 \frac{R_Г + r_б}{R_Г + r_б + r_з} = \alpha I_3 \gamma_3,$$

где

$$\gamma_3 = \frac{R_Г + r_б}{R_Г + r_б + r_з}$$

— коэффициент *тока распределения*, показывающий, какая доля коллекторного тока ответвляется в цепь эмиттера.

¹ Генератор электродвижущей силы $\bar{e}_{т.г.}$, как и другие генераторы шумов, не имеет определенной полярности, поэтому принятая на рис. 4-30, а полярность является условной. При обратной полярности изменятся направления токов I_3 и αI_3 по сравнению с положительными направлениями, принятыми на рис. 4-30, а. Это никак не отразится на значении коэффициента шума, так как он определяется квадратами токов.

Складывая обе составляющие I'_3 и I''_3 , получаем уравнение:

$$I_3 = \frac{\bar{e}_{т.г}}{R_r + r_6 + r_3} + \alpha I_3 \gamma_3,$$

из которого легко определяются токи I_3 и αI_3 . Последний полностью протекает в цепи нагрузки и, следовательно, является искомым компонентом $\bar{i}_4$.

Запишем ток $\bar{i}_1^2$ в следующем виде:

$$\bar{i}_1^2 = \left(\frac{\alpha}{1 - \alpha \gamma_3} \right)^2 \frac{\bar{e}_{т.г}^2}{(R_r + r_6 + r_3)^2}. \quad (4-111a)$$

Поскольку э. д. с. $\bar{e}_{т.б}$ в схеме на рис. 4-29 включена последовательно с э. д. с. $\bar{e}_{т.г}$, компонент $\bar{i}_2$ будет иметь такую же структуру, как (4-111a):

$$\bar{i}_2^2 = \left(\frac{\alpha}{1 - \alpha \gamma_3} \right)^2 \frac{\bar{e}_{т.б}^2}{(R_r + r_6 + r_3)^2}. \quad (4-111б)$$

Для определения компонента $\bar{i}_3$, обусловленного источником тока $\bar{i}_{3б'}$, используем эквивалентную схему на рис. 4-30, б. Из нее методом суперпозиций (почередно обрывая цепи αI_3 и $\bar{i}_{3б'}$) получаем:

$$\bar{i}_3^2 = \left(\frac{\alpha \gamma_3}{1 - \alpha \gamma_3} \right)^2 \bar{i}_{3б'}^2. \quad (4-111в)$$

Компонент $\bar{i}_4$, обусловленный источником тока $\bar{i}_{6'к}$, определяется из эквивалентной схемы на рис. 4-30, в и равен сумме токов $\bar{i}_{6'к} + \alpha I_3$, где ток αI_3 рассчитывается методом суперпозиции. В результате находим:

$$\bar{i}_4^2 = \frac{\bar{i}_{6'к}^2}{(1 - \alpha \gamma_3)^2}. \quad (4-111г)$$

Наконец, ток $\bar{i}_{5к}$ полностью протекает в цепи нагрузки, не ответвляясь в цепь эмиттера, поэтому

$$\bar{i}_5^2 = \bar{i}_{5к}^2. \quad (4-111д)$$

Подставив полученные компоненты в (4-110) и используя формулы (4-108), нетрудно найти коэффициент шума схемы. Чтобы избежать громоздкого выражения, которое получается в общем случае, положим $r_3 \ll R_r + r_6$ (т. е. $\gamma_3 = 1$) и $\alpha = 1$ (за исключением тех случаев, когда имеется двучлен $1 - \alpha$). Оба предположения вполне оправдываются на практике. Тогда после соответствующих преобразований и учета (1-3) получим коэффициент шума в следующем виде:

$$F_{ш} \approx 1 + \frac{r_6}{R_r} + \frac{(R_r + r_6)^2}{2\varphi_r R_r} \left[I_{к0} + (1 - \alpha) I_3 + \frac{A}{2qf} \right], \quad (4-112)$$

где $A = A_3 + A_к$.

Можно показать, что значение коэффициента шума почти не зависит от схемы включения транзистора, поэтому последующие выводы носят достаточно общий характер.

Анализ коэффициента шума. Из формулы (4-112) легко заметить, что коэффициент шума имеет минимальное значение при определенном сопротивлении источника сигнала R_r . Дифференцируя

(4-112) по R_r и приравняв производную нулю, легко найти значение оптимального сопротивления:

$$R_{r, \text{опт}} = r_6 \sqrt{1 + \frac{2\varphi_T/r_6}{I_{k0} + (1-\alpha)I_s + A/2qf}}. \quad (4-113a)$$

Если соблюдается условие $I_{k0} + A/2qf \ll (1-\alpha)I_s$, то с учетом выражения (4-22) сопротивление $R_{r, \text{опт}}$ можно записать в следующей форме:

$$R_{r, \text{опт}} \approx r_6 \sqrt{1 + \frac{2r_s}{(1-\alpha)r_6}}. \quad (4-113b)$$

Например, если $r_6 = 200 \text{ Ом}$; $r_s = 25 \text{ Ом}$ и $\alpha = 0,98$, то $R_{r, \text{опт}} \approx 700 \text{ Ом}$. Минимум функции $F_{\text{ш}}(R_r)$ выражен неявно, поэтому увеличение и уменьшение сопротивления R_r в 2—3 раза по сравнению с $R_{r, \text{опт}}$ изменяют $F_{\text{ш}}$ всего лишь на 25—30%.

Выражение (4-112) показывает, что коэффициент шума зависит от диапазона частот: в области низких частот возрастает слагаемое $A/2qf$, а в области высоких частот — слагаемое $I_s(1-\alpha)$ (поскольку коэффициент α уменьшается с ростом частоты). Соответствующая частотная характеристика показана на рис. 4-31. При анализе шумов областью средних частот является такой диапазон, в котором избыточные шумы уже несущественны, а заметный спад α еще не наступил. Этот диапазон обычно лежит в пределах от нескольких килогерц до сотен килогерц и выше. Оценим коэффициент шума в области средних частот, пренебрегая в формуле (4-112) членом $A/2qf$ и подставляя $I_s = 1 \text{ мА}$; $I_{k0} = 10 \text{ мкА}$; $\alpha = 0,98$; $r_6 = 200 \text{ Ом}$ и найденное выше значение $R_r = 700 \text{ Ом}$. Тогда получается характерное значение $F_{\text{ш}} \approx 2$ (т. е. 3 дБ).

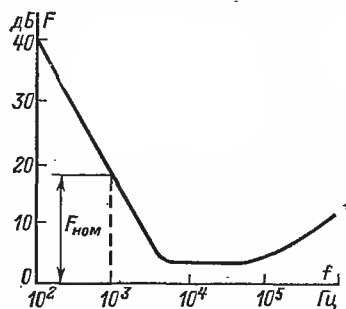


Рис. 4-31. Частотная характеристика коэффициента шума в логарифмическом масштабе.

В области высоких частот в формуле (4-112) следует пренебречь членом $A/2qf$, а в качестве коэффициента α подставить модуль, соответствующий данной частоте. Например, если сохранить использованные ранее значения параметров и уменьшить α до 0,67 (это примерно соответствует частоте f_a), то коэффициент шума будет иметь значение $F_{\text{ш}} = 6,8$ (т. е. 8,3 дБ). С физической точки зрения возрастание шумов в области высоких частот объясняется резким увеличением базового тока $I_6 = I_s(1-\alpha)$ и связанных с ним рекомбинационных флуктуаций.

В области низких частот коэффициент шума быстро нарастает, и практически в формуле (4-112) можно пренебречь всеми членами,

кроме $A/2qf$. Тогда получим для этой области:

$$F_{\text{ш}} \approx \frac{(R_{\Gamma} + r_6)^2}{4q\varphi_{\Gamma} R_{\Gamma}} \frac{A}{f} = \frac{A^*}{f}. \quad (4-114)$$

Как видим, параметр A^* имеет сложную структуру и, в частности, существенно зависит от значения сопротивления R_{Γ} . Оптимальным значением в данном случае является $R_{\Gamma} = r_6$; оно меньше оптимального значения (4-113), найденного для широкого диапазона частот. Поскольку коэффициенты $A_{\text{в}}$ и $A_{\text{к}}$, входящие в состав параметра A^* , рассчитать трудно, величину A^* определяют экспериментально путем измерения коэффициента шума при определенном режиме транзистора, определенном значении R_{Γ} (обычно 0,5—1 кОм) и определенной частоте (как правило, 1 кГц). Типичные значения коэффициента шума, измеренного таким путем лежат в пределах 10—20 дБ ($F_{\text{ш}} = 10 \div 100$) для обычных транзисторов и 5—10 дБ ($F_{\text{ш}} = 4 \div 10$) для специальных «малошумящих» транзисторов [12, 66]. Соответствующие значения A^* составляют 10^4 — 10^5 и $4 \cdot 10^3$ — 10^4 с⁻¹. Отсюда легко оценить типичные значения коэффициента A в формуле (4-114). Полагая $R_{\Gamma} = 500$ Ом (т. е. $R_{\Gamma} \gg r_6$), получаем $A = 10^{-18} \div 10^{-16}$ А²; меньшее значение соответствует «малошумящим» транзисторам.

Зависимость низкочастотных шумов от режима определяется приближенными выражениями для коэффициентов $A_{\text{в}}$ и $A_{\text{к}}$, которые указаны в формулах (4-107). Однако эти зависимости носят качественный характер. Для количественных расчетов следует пользоваться типовыми или индивидуальными кривыми, снятыми экспериментально.

Из формулы (4-112) видно, что при данной частоте коэффициент шума увеличивается с ростом рабочего тока (через член $I_{\text{в}}(1 - \alpha)$), а также через коэффициент A и температуры (через член $I_{\text{к0}}$). Однако для расчетов этих зависимостей также рекомендуется использовать опытные данные.

Мощность и напряжение шума. Приведем соотношения, связывающие величину коэффициента шума с мощностью и напряжением шумов на входе каскада [72]. Если в правой части формулы (4-109) разделить числитель и знаменатель на коэффициент усиления мощности, то в числителе будет стоять полная мощность шумов, приведенная ко входу, а в знаменателе — мощность тепловых шумов в сопротивлении R_{Γ} . Последняя, как известно, равна:

$$P_{\Gamma} = 4kT\Delta f.$$

Поэтому приведенная (ко входу) мощность шумов выражается следующим образом:

$$P_{\text{ш}} = 4kTF_{\text{ш}}\Delta f \approx 1,6 \cdot 10^{-20} F_{\text{ш}}\Delta f. \quad (4-115a)$$

Например, если $F_{\text{ш}} = 100$ и $\Delta f = 2$ кГц, то $P_{\text{ш}} = 3 \cdot 10^{-15}$ Вт. Формула (4-115a) действительна только для «белого шума» или при условии $\Delta f \ll f$; в противном случае нужно определять мощность путем интегрирования по диапазону частот. Так, в области низких частот, где спектр шумов имеет вид (4-114), получается выражение

$$P_{\text{ш}} = 4kTA^* \ln \frac{f_2}{f_1} = 3,6 \cdot 10^{-20} A^* \lg \frac{f_2}{f_1}, \quad (4-115b)$$

где f_2 и f_1 — граничные частоты. Пусть, например, $A^* = 10^5$ с⁻¹; $f_2 = 2$ кГц; $f_1 = 50$ Гц; тогда $P_{\text{ш}} \approx 6 \cdot 10^{-16}$ Вт.

Приведенное (ко входу) напряжение шумов получается из соотношения $U_{\text{ш}}/R_{\Gamma} = P_{\text{ш}}$ и для вариантов (4-115а) и (4-115б) имеет вид:

$$\bar{U}_{\text{ш}} = \sqrt{4kTR_{\Gamma}F_{\text{ш}}\Delta f} \approx 1,26 \cdot 10^{-10} \sqrt{R_{\Gamma}F_{\text{ш}}\Delta f}; \quad (4-116a)$$

$$\bar{U}_{\text{ш}} = \sqrt{4kTR_{\Gamma}A^* \ln \frac{f_2}{f_1}} \approx 2 \cdot 10^{-10} \sqrt{R_{\Gamma}A^* \lg \frac{f_2}{f_1}}. \quad (4-116б)$$

Для тех же параметров, что и в предыдущих примерах, и при $R_{\Gamma} = 500 \text{ Ом}$ получим соответственно $\bar{U}_{\text{ш}} \approx 1,3 \text{ мкВ}$ и $\bar{U}_{\text{ш}} \approx 1,8 \text{ мкВ}$.

4-10. СОСТАВНЫЕ ТРАНЗИСТОРЫ

Для того чтобы повысить значение коэффициента усиления β , нужно согласно (4-75) уменьшить толщину базы, что, конечно, представляет трудности технологического характера. Между тем, соединяя определенным образом два транзистора, можно резко повысить коэффициент усиления такой комбинации, рассматриваемой

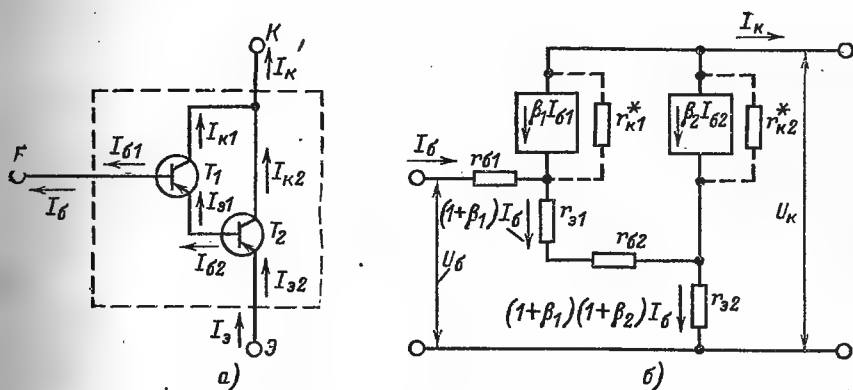


Рис. 4-32. Составной транзистор (а) и его эквивалентная схема (б).

как единое целое. Такая комбинация (часто выполняемая на одной пластине с внутренними соединениями и тремя внешними выводами) называется *составным транзистором* или *схемой Дарлингтона* (рис. 4-32, а).

Покажем, что составной транзистор действительно имеет коэффициент β , значительно больший, чем у обоих его компонентов. Задавая приращение тока $dI_{\text{б}} = dI_{\text{б}1}$, получаем:

$$dI_{\text{э}1} = (1 + \beta_1) dI_{\text{б}} = dI_{\text{б}2};$$

$$dI_{\text{к}} = dI_{\text{к}1} + dI_{\text{к}2} = \beta_1 dI_{\text{б}} + \beta_2 [(1 + \beta_1) dI_{\text{б}}].$$

Деля $dI_{\text{к}}$ на $dI_{\text{б}}$, находим результирующий дифференциальный коэффициент передачи:

$$\beta_{\Sigma} = \beta_1 + \beta_2 + \beta_1\beta_2. \quad (4-117a)$$

Поскольку всегда $\beta \gg 1$, можно считать:

$$\beta_{\Sigma} \approx \beta_1 \beta_2. \quad (4-1176)$$

Значение β_{Σ} может составлять несколько тысяч при использовании рядовых транзисторов.

Следует подчеркнуть, что различие коэффициентов β_1 и β_2 может иметь место даже в случае одинаковых транзисторов, поскольку ток $I_{\beta 2}$ в $1 + \beta_2$ раз больше тока $I_{\beta 1}$ (это вытекает из очевидного равенства $I_{\beta 2} = I_{\beta 1}$). Различие в токах сказывается также на величинах сопротивлений r_{β} и r_k [см. (4-22) и (4-24)]. Токи обоих транзисторов можно выровнять, задавая в узловую точку $\mathcal{E}_1 - B_2$ ток, близкий к $I_{\beta 1}$ (см. рис. 9-7). Тогда параметры обоих транзисторов, в том числе коэффициенты β , будут почти одинаковыми.

Воспользовавшись эквивалентной схемой на рис. 4-32, б, рассмотрим сопротивления эмиттерного и коллекторного переходов составного транзистора. Пренебрегая сначала сопротивлениями r_k^* и считая $r_{\beta \Sigma} = r_{\beta 1}$, определим результирующее сопротивление $r_{\beta \Sigma}$ из соотношения (4-103а). Для этого нужно предварительно найти параметр h_{11} , т. е. входное сопротивление составного транзистора при коротком замыкании на выходе. После этого, учитывая (4-117а), легко получить:

$$r_{\beta \Sigma} = r_{\beta 2} + \frac{r_{\beta 1} + r_{\beta 2}}{1 + \beta_2} \approx 2r_{\beta 2}. \quad (4-118а)$$

Результирующее сопротивление коллектора $r_{k\Sigma}^*$ определяется из соотношения (4-103г) как $1/h_{22}$, где h_{22} — выходная проводимость при холостом ходе на входе, т. е. при $I_6 = 0$ (это означает, что в схеме на рис. 4-32, б отсутствует генератор $\beta_1 I_6$). Полагая для простоты $r_{\beta 2} = 0$ (что не приводит к заметной ошибке), легко найти токи в трех оставшихся ветвях схемы при подаче напряжения U_k . Складывая эти токи и пренебрегая суммой $r_{\beta 1} + r_{\beta 2}$ по сравнению с r_{k1}^* , приходим к следующему выражению¹:

$$r_{k\Sigma}^* = r_{k2}^* \parallel \left(\frac{r_{k1}^*}{1 + \beta_2} \right) \approx \frac{r_{k2}^*}{2}. \quad (4-118б)$$

Приближения в формулах (4-118) основаны на отмеченном выше различии токов $I_{\beta 1}$ и $I_{\beta 2}$. В случае одинаковых токов можно считать $r_{\beta \Sigma} \approx r_{\beta 2}$ и $r_{k\Sigma}^* \approx r_{k2}^*/(1 + \beta_2)$.

Наконец, оценим результирующий тепловой ток составного транзистора. Введя генераторы I_{k01}^* и I_{k02}^* в схему на рис. 4-32, б, пренебрегая сопротивлениями r_k и полагая $I_6 = 0$ (а следовательно, и $\beta_1 I_6 = 0$), легко убедиться, что коллекторный ток будет состоять из трех компонентов: I_{k01}^* , I_{k02}^* и $\beta_2 I_{k01}^*$; тогда

$$I_{k0\Sigma}^* = I_{k02}^* + (1 + \beta_2) I_{k01}^*. \quad (4-118в)$$

Ясно, что тепловой ток в составном транзисторе больше, чем в каждом из компонентов. Этот серьезный недостаток затрудняет

¹ Здесь и далее символ $\parallel$ означает параллельное соединение соответствующих сопротивлений.

использование германиевых составных транзисторов при повышенных температурах. В случае кремниевых транзисторов такие затруднения не возникают.

Число компонентов составного транзистора может быть равно не только двум, но и трем. Более подробно составные транзисторы рассмотрены в работе [73].

4-11. ДОПУСТИМАЯ МОЩНОСТЬ И ОСОБЕННОСТИ МОЩНЫХ ТРАНЗИСТОРОВ

Поскольку в активном режиме токи I_b и I_k почти одинаковы, а напряжение U_k значительно больше, чем U_b , то основная часть мощности потерь выделяется в области коллекторного перехода. Каждый транзистор характеризуется предельно допустимой температурой перехода, при превышении которой параметры резко ухудшаются. Исходя из соотношения (2-41), легко прийти к следующей зависимости между допустимой мощностью рассеяния, допустимой температурой перехода и температурой окружающей среды:

$$P_{\text{доп}} = \frac{T_{\text{пер. доп}} - T_{\text{окр}}}{R_t}, \quad (4-119)$$

где R_t — тепловое сопротивление переход—среда, которое, как и величина $T_{\text{пер. доп}}$, указывается в справочниках. Типичным значением R_t для маломощных транзисторов является 0,5—0,7 °C/мВт. Для мощных транзисторов это значение в десятки раз меньше. Типичными значениями $T_{\text{пер. доп}}$ являются 150—200 °C для кремния и 90—100° C для германия.

Из формулы (4-119) следует, что допустимая мощность уменьшается с ростом окружающей температуры и что главным путем повышения мощности является уменьшение сопротивления R_t , т. е. улучшение теплоотвода.

Мощные транзисторы характерны большими рабочими токами и соответственно большими площадями p - n переходов (до 1 см²). И тот, и другой фактор отражаются на значении основных параметров и придают мощным транзисторам определенную специфику.

Так, при большой площади переходов трудно реализовать тонкую базу, особенно сплавным методом, а это согласно (4-54) приводит к сравнительно низкой предельной частоте f_t . У мощных $сплавных$ транзисторов частота f_t не превышает 100—200 кГц, но даже у дрейфовых мощных транзисторов, у которых частота f_t достигает 50—100 МГц, она все же ниже, чем у маломощных дрейфовых транзисторов, у которых в настоящее время достигнуты значения f_t , составляющие несколько гигагерц. Кроме того, коллекторная емкость у мощных транзисторов может составлять сотни пикофард, так что в целом мощные транзисторы являются сравнительно низкочастотными.

Большие рабочие токи приводят к резкому уменьшению сопротивлений r_b и r_k . Из выражения (4-22) следует, что при токе больше

100 мА сопротивление эмиттерного перехода ничтожно мало и с ним практически можно не считаться. Поэтому в схеме ОЭ входным сопротивлением будет по существу только сопротивление базы r_6 (см. рис. 4-24) *. Последнее при высоких уровнях инжекции модулируется (см. рис. 2-34) и обычно лежит в пределах до 10 Ом. Малое значение входного сопротивления не является препятствием для применения мощных транзисторов, если связь с источником сигнала осуществляется через трансформатор. Сопротивление коллекторного перехода при токах около 1 А составляет всего несколько килоом, а сопротивление r_k^* в схеме ОЭ — сотни ом.

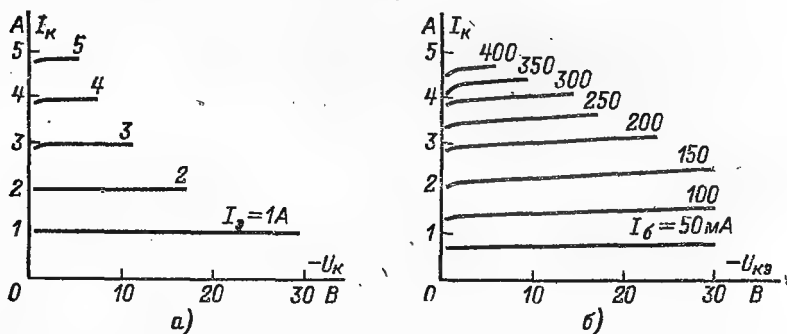


Рис. 4-33. Коллекторные характеристики мощного транзистора.

а — в схеме ОБ; б — в схеме ОЭ.

Тепловой ток коллектора, пропорциональный площади перехода, доходит у мощных транзисторов до десятков миллиампер. С приближением напряжения к максимально допустимому значению тепловой ток увеличивается в несколько раз в связи с возрастающей ролью термогенерации и ударной ионизации в переходе, а также саморазогрева.

Как известно, при больших эмиттерных токах наблюдается уменьшение коэффициентов передачи α и β (см. рис. 4-19, б). У мощных транзисторов эта зависимость весьма существенна: коэффициент β в рабочем диапазоне токов может уменьшиться в 2—3 раза, а иногда и больше. В связи с этим характеристики коллекторного семейства заметно сближаются в области больших токов, особенно в схеме ОЭ (рис. 4-33). Результатом такой неоднородности харак-

* Такие мощные транзисторы, у которых $r_3 = 0$, удобно характеризовать крутизной S в схеме ОЭ. Используя условие $r_3 = 0$, получаем:

$$dI_6 = dU_6/r_6 \text{ и } dI_k \approx \beta dI_6;$$

тогда

$$S = \frac{\partial I_k}{\partial U_6} = \frac{\beta}{r_6}.$$

В зависимости от тока крутизна лежит в пределах от 1—2 до 10—20 А/В.

геристик являются, конечно, нелинейные искажения при усилении сигналов.

Одной из главных проблем при конструировании и эксплуатации мощных транзисторов является обеспечение теплоотвода [60]. Эта проблема решается путем непосредственного контакта коллектора с корпусом транзистора и применением внешних радиаторов. В некоторых случаях радиатором может служить шасси. Если схема усилителя не допускает электрического контакта между коллектором и шасси (т. е. «заземления» коллектора), можно использовать тонкие слюдяные прокладки, изолирующие корпус транзистора от радиатора или шасси. Такие прокладки сравнительно слабо влияют на тепловой контакт, т. е. не сильно увеличивают тепловое сопротивление системы R_t . Тепловое сопротивление убывает с увеличением площади теплоотвода и потому имеет минимальное значение у наиболее мощных транзисторов. Тепловое сопротивление в отсутствие радиатора не допускает рассеяния мощности больше 1—4 Вт. Наличие радиатора снижает значение R_t и приближает его к тепловому сопротивлению участка переход—корпус. Соответственно допустимая мощность сильно возрастает. На рис. 4-34 показана типичная кривая, иллюстрирующая роль радиатора.

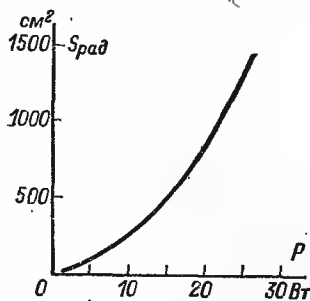


Рис. 4-34. Зависимость допустимой мощности рассеяния от площади радиатора.

В заключение рассмотрим одно явление, которое характерно для мощных транзисторов, но, вообще говоря, имеет место во всех случаях, когда падение напряжения $I_6 r_{61}$ в активной области базы (область 1 на рис. 4-14) примерно равно φ_T и более. Это явление, получившее название *эффект оттеснения тока* [74, 58], состоит в следующем.

Поскольку ток базы протекает параллельно плоскостям эмиттера и коллектора, потенциал базы уменьшается от центра эмиттера к его периферийным частям (рис. 4-35, а). Соответственно разность потенциалов $U_{эб}$ будет минимальна в центре и максимальна на краях эмиттера. В значительно большей степени будут различаться плотности тока, так как они согласно (2-32) экспоненциально зависят от $U_{эб}$.

Если падение потенциала вдоль базы превышает $2\varphi_T$, то плотности тока в центре и на периферии эмиттера различаются на порядок и более, т. е. фактически весь ток эмиттера обусловлен его периферийной частью (на рис. 4-35, а она условно зачернена), а центральная часть почти «бездействует». Значит, плотность тока эмиттера, рассчитанная без учета эффекта оттеснения (т. е. исходя из минимального значения $U_{эб}$ в центре эмиттера), будет намного меньше ее реальных значений. Соответственно эффект оттеснения «форсирует» те явления, которые связаны с высоким

уровнем инжекции: все они [в частности, зависимость $\beta(I_E)$] наступают при заметно меньших токах, чем без учета этого эффекта.

Для того чтобы избежать эффекта оттеснения или уменьшить его влияние, следует уменьшать размеры эмиттера настолько, чтобы падение напряжения $I_E r_{E1}$ под эмиттером было меньше φ_T . Можно, например, выполнять эмиттер не в виде диска (рис. 4-3 и 4-14), а в виде кольца, сечение которого соответствует зачерченному участку на рис. 4-35, а. В случае планарных транзисторов используют «многополосковые» гребенчатые структуры (рис. 4-35, б), у которых ширина каждой полоски достаточно мала (т. е. $I_E r_{E1} < \varphi_T$), а суммарная ширина достаточно велика, чтобы обеспечить приемлемую плотность тока.

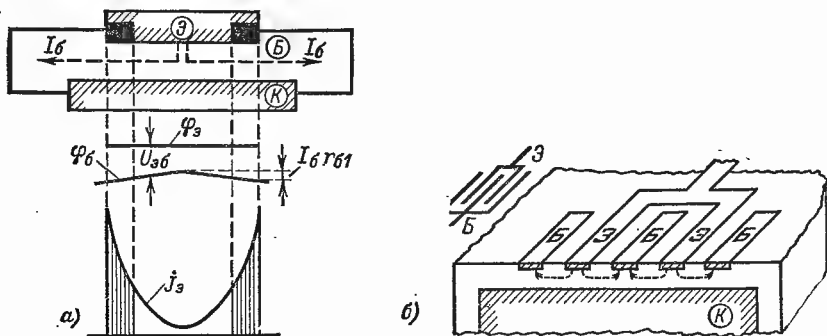


Рис. 4-35. Эффект оттеснения тока в эмиттере.

а — разрез дисковой структуры, распределение потенциала базы и плотности эмиттерного тока; б — планарная структура с многополосковыми (гребенчатыми) эмиттером и базой.

У маломощных транзисторов эффект оттеснения тока выражен слабо, так как, во-первых, малы размеры эмиттера (от которых зависит сопротивление r_{E1}) и, во-вторых, мал рабочий ток I_K , а значит, и ток I_E .

4-12. ДРЕЙФОВЫЕ ТРАНЗИСТОРЫ

Как отмечалось в § 4-1, в неоднородной базе имеется собственное электрическое поле независимо от уровня инжекции. Поэтому механизм движения инжектированных носителей у транзисторов с неоднородной базой преимущественно дрейфовый, откуда следует их общее название — *дрейфовые транзисторы* [14, 75]. Поскольку, однако, неоднородность базы достигается путем диффузии примесного материала, можно встретиться с термином «транзистор с диффузионной базой», который отражает технологию изготовления, а не механизм движения инжектированных носителей. В последние годы дрейфовые транзисторы стали играть главную роль в связи с развитием микроэлектроники, которой свойственна планарная диффузионная технология (см. § 4-13).

Особенности дрейфовых транзисторов. На рис. 4-36 изображены две структуры дрейфового транзистора, различие между которыми обусловлено технологическими особенностями.

Сравнивая структуры дрейфового и бездрейфового транзисторов (см. рис. 4-3), видим, что они заметно различаются. А именно, дрейфовый транзистор значительно более несимметричен: коллекторный слой много толще двух других слоев, велика разница в площадях эмиттера и коллектора, по существу отсутствует пассивная область базы. Все эти особенности обусловлены тем, что в случае дрейфовых транзисторов исходная пластина полупроводника служит основной для создания слоев базы и эмиттера, а сама она в дальнейшем играет роль к о л л е к т о р а, тогда как в случае бездрейфовых (сплавных) транзисторов исходная пластина служит основной для создания слоев эмиттера и коллектора, а сама она в дальнейшем играет роль б а з ы.

Резкая асимметрия дрейфовых транзисторов¹ делает их практически н е о б р а т и м ы м и приборами. Кроме того, она осложняет анализ тех случаев, когда существенна н е о д н о м е р н о с т ь транзистора. Однако при анализе нормального усилительного режима использование

одномерной модели вполне оправдано, поскольку движение инжектированных носителей в активной области базы происходит почти строго по оси x (рис. 4-36).

Как уже отмечалось, база дрейфовых транзисторов получается путем диффузии примесей в глубь исходной пластины (на рис. 4-36 вдоль оси x).

Согласно законам диффузии распределение примесных атомов в слое базы должно выражаться дополнительной функцией ошибок [см. (4-149)]:

$$N_6(x) = N_6(0) \operatorname{erf} \left(\frac{x}{L_N} \right),$$

где x — расстояние, отсчитываемое от эмиттера в глубь базы; L_N — средняя длина диффузии п р и м е с и за время нагрева.

Поскольку функция $\operatorname{erf}(z)$ близка к спадающей экспоненте (см. сноску на с. 84), можно записать приближенное выражение

$$N_6(x) \approx N_6(0) e^{-\frac{x}{L_N}}, \quad (4-120)$$

которым широко пользуются при анализе дрейфовых транзисторов.

¹ Помимо геометрической асимметрии, отмеченной выше, дрейфовым транзисторам свойственна асимметрия электрическая: коллекторный и эмиттерный слои в отличие от сплавных транзисторов имеют резко отличающиеся удельные сопротивления.

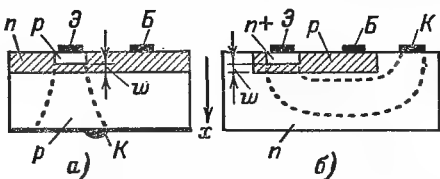


Рис. 4-36. Упрощенные структуры дрейфовых транзисторов.

а — выводы электродов в разных плоскостях (мезотехнология и сплавно-диффузионная технология); б — выводы электродов в одной плоскости (планарная технология). Точками показаны линии тока от эмиттера к коллектору.

В работе [76] на основании измерений отмечается, что в реальных приборах распределение примесей ближе к экспоненте, чем к функции ошибок. Поэтому функция (4-120) является не только приближением, но и экспериментально установленным фактом.

Диффузионная технология позволяет получить очень тонкую базу, что само по себе (даже без учета распределения примесей) приводит к ряду важных следствий. А именно, при прочих равных условиях существенно уменьшается время диффузии t_D и увеличивается коэффициент передачи β , поскольку оба эти параметра зависят от квадрата толщины базы¹. Толщина базы у дрейфовых транзисторов в 10—20 раз меньше, чем у сплавных, а потому время диффузии t_D и постоянная времени τ_a оказываются меньше в сотни раз; соответственно увеличивается граничная частота f_a . Коэффициент передачи β по тем же соображениям должен был бы доходить до 3000 и больше. На самом деле он значительно меньше и обычно не превышает 200—300. Это объясняется тем, что величины α и β зависят не только от толщины базы, но также от времени жизни и коэффициента инжекции. В связи с повышенной концентрацией примесей вблизи эмиттера время жизни в базе дрейфового транзистора значительно меньше, чем у сплавного (см. рис. 1-24, а), а коэффициент инжекции более заметно отличается от единицы [см. (2-35)]. Тем не менее у специальных типов дрейфовых транзисторов [77] удается получить значения β до 5000 и более, уменьшая толщину базы до долей микрона². Однако такие значения β получаются за счет резкого уменьшения рабочих напряжений: напряжение смыкания (4-78) составляет у этих транзисторов всего 1—1,5 В.

Теперь учтем неравномерное распределение примесей в базе на примере p - n - p транзистора (рис. 4-37, где L_d — длина диффузии доноров) и покажем те следствия, к которым приводит такая неравномерность. Прежде всего очевидно, что слой базы, прилегающий к коллекторному переходу, является почти собственным полупроводником, так как здесь продифундировавшие донорные атомы в значительной мере компенсируют акцепторные атомы исходного кристалла. Следовательно, удельное сопротивление этого слоя базы велико и коллекторный переход оказывается довольно широким. Соответственно емкость C_k получается значительно (почти на порядок) меньшей, чем у сплавных транзисторов; она может составлять несколько пикофард и менее. По вполне понятным причинам коллекторный переход является плавным, а не ступенчатым, и потому емкость C_k описывается формулой (2-83).

По мере удаления от коллектора в глубь базы концентрация доноров растет, а удельное сопротивление уменьшается. Сопротивление активной области базы можно рассматривать как результат

¹ Зависимость $t_D \sim \omega^2$ очевидна из (2-88), а зависимость $\beta \sim 1/\omega^2$ — из (4-75б).

² Транзисторы этого типа принято называть «супербета-транзисторами». Для них характерны значения $\omega = 0,2 \div 0,3$ мкм.

параллельной соединяющей отдельные слои, имеющих разную удельную проводимость. Очевидно, что сопротивление r_6 определяется в основном тем участком базы, который прилегает к эмиттерному переходу и имеет наибольшую удельную проводимость. Поэтому, несмотря на значительно меньшую толщину базы w , значение r_6 у дрейфовых транзисторов примерно такое же, как у сплавных, и даже меньше.

Эмиттерный переход у дрейфовых транзисторов, как правило, ступенчатый. Поскольку граничная концентрация доноров в базе $N_d(0)$ велика, концентрация акцепторов в эмиттере должна быть еще большей и эмиттерный переход получается очень узким. В результате при подаче на эмиттер отрицательного запирающего напряжения этот переход легко пробивается. Пробой носит туннельный характер (см. § 2-7) и происходит при очень небольшом напряжении (1—2 В). Пробой эмиттера оказывает значительное влияние на работу многих импульсных схем, в которых запираение транзистора является необходимым элементом рабочего цикла. Эта важная специфика дрейфовых транзисторов не является, однако, препятствием для применения их в ключевых схемах, так как пробой перехода при ограниченном токе является обратимым явлением (как в полупроводниковом стабилитроне) и не представляет никакой опасности. Инжекция в режиме пробоя, как известно, отсутствует, и, следовательно, по коллекторной цепи транзистор остается запертым.

Меньшая ширина эмиттерного перехода у дрейфовых транзисторов при прочих равных условиях означает большое значение барьерной емкости C_b . Это обстоятельство вместе с гораздо более высокой частотой f_a делает существенным влияние емкости C_b на коэффициент инжекции [см. (4-30)]. Иначе говоря, частотные свойства дрейфовых транзисторов могут ограничиваться не временем пролета, а постоянной времени τ_v . Для того чтобы уменьшить влияние емкости C_b , дрейфовые транзисторы часто используют при большем токе эмиттера, например 4—5 мА вместо 1 мА. Тогда сопротивление r_6 уменьшается и постоянная времени $r_6 C_b$ оказывается достаточно малой.

Конечно, значения барьерных емкостей C_k и C_b зависят не только от соотношения удельных сопротивлений слоев, но и от площадей

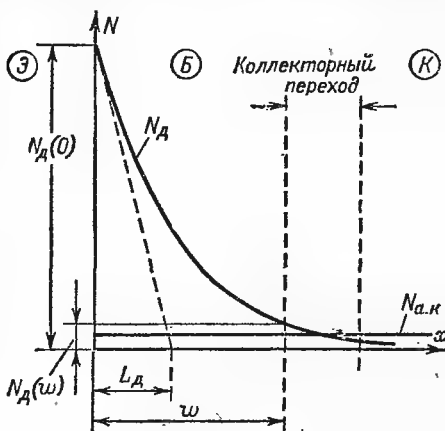


Рис. 4-37. Распределение примесей в базе дрейфового транзистора.

переходов. Технология изготовления дрейфовых транзисторов обес­печивает гораздо меньшие площади, чем сплавная технология, и это способствует уменьшению барьерных емкостей. Однако относитель­ная роль коллекторной емкости у дрейфовых транзисторов выше, поскольку в связи со специфической их конфигурацией (см. рис. 4-36) площадь коллектора существенно больше площади эмиттера.

Заметим еще, что коллекторный слой у дрейфовых транзисторов может иметь сравнительно большое сопротивление. Это объясняется, во-первых, значительной толщиной коллектора (она близка к тол­щине исходной пластинки на рис. 4-36) и, во-вторых, тем, что исход­ная пластинка имеет довольно большое удельное сопротивление (в противном случае было бы трудно обеспечить высокое пробив­ное напряжение, малую коллекторную емкость, а также достаточный перепад концентрации примеси в базе, поскольку концентрация $N_d(x)$ растет вместе с концентрацией $N_{a,k}$, см. рис. 4-37). Большое сопротивление коллекторного слоя особенно нежелательно в клю­чевых схемах¹.

Распределение носителей в базе. Анализ дрейфового тран­зистора сложнее, чем бездрейфового. В самом деле, при наличии поля нельзя пользоваться диффузионными уравнениями (1-79), а нужно использовать полные уравнения непрерывности (1-78) вместе с уравнением Пуассона (1-80). Кроме того, следует учиты­вать, что подвижности носителей, а значит, и коэффициенты диф­фузии являются функциями концентрации примесей и, следова­тельно, существенно меняются вдоль базы. При этом получается система нелинейных уравнений, разрешить которую в общем виде не удастся. Поэтому подвижность так или иначе усредняют и считают постоян­ной. Считают также постоянным собственное поле в базе, причем это приближение имеет под собой довольно прочное основание, если положить в основу распределение примеси (4-120). Действительно, в § 1-12 для такого распределения было получено [см. (1-94)]:

$$E = \frac{\varphi_T}{L_d}. \quad (4-121)$$

Следовательно, потенциал вдоль базы меняется линейно:

$$\varphi_E(x) = \varphi_E(0) + \varphi_T \frac{x}{L_d},$$

где $\varphi_E(0)$ — электростатический потенциал на эмиттерной гра­нице базы.

Соответствующая зонная диаграмма показана на рис. 4-38. Поскольку эта диаграмма, как и прежние, построена применительно

¹ В последние годы найдены эффективные методы, позволяющие обеспечить малое сопротивление коллекторного слоя при достаточно высоком его удель­ ном сопротивлении (см. § 4-13).

к электронам, положительный градиент потенциала соответствует положительному полю, направленному по оси x . Такое поле способствует движению инжектированных дырок¹.

В равновесном состоянии электроны в базе дрейфового транзистора p - n - p распределены почти так же, как доноры, т. е. их концентрация экспоненциально уменьшается от эмиттера к коллектору. Концентрация дырок согласно (1-16) должна при этом экспоненциально возрастать (рис. 4-39). Эти распределения поддерживаются благодаря тому, что диффузионные составляющие токов уравниваются дрейфовыми составляющими (больцмановское равновесие).

При включении достаточно большого отрицательного смещения на коллектор концентрация $p(\omega)$ согласно (2-13а) падает почти до нуля. Так как в точке $x = 0$ концентрация дырок с самого начала была очень малой, приходим к выводу, что распределение $p(x)$, определяющее ток $I_{к0}$, будет почти таким же, как и в бездрейфовых транзисторах [см. (4-28)]. Однако тепловой ток дрейфовых транзисторов имеет меньшее значение из-за малой площади и меньшей толщины базы [см. (4-29б)].

Когда задан ток эмиттера и имеет место инжекция, распределение носителей получается весьма своеобразным, отнюдь не линейным, как у бездрейфовых транзисторов. В самом деле, если бы движение дырок было чисто дрейфовым на всем протяжении базы, концентрация p не должна была бы зависеть от координаты x , т. е. представлялась бы горизонтальной линией. Однако вблизи точки $x = \omega$, где $p(\omega) = 0$, неизбежен спад концентрации, а следовательно, возрастает роль диффузии. Поэтому в относительном масштабе распределение $p(x)$ имеет приблизительно такой вид, как показано на рис. 4-40 (кривые $\eta = 1, 2, 4$).

Для того чтобы найти распределения, показанные на рис. 4-40, подставим в уравнение непрерывности (1-78а) напряженность поля из (4-121); при этом выпадает член с dE/dx , поскольку $E = \text{const}$. Кроме того, учитывая замечания, сделанные применительно к распределению (4-17а), пренебрежем с самого начала генерационным членом p_0/τ . Тогда стационарное уравнение непрерывности для

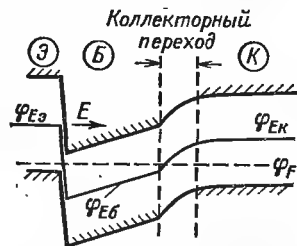


Рис. 4-38. Зонная диаграмма дрейфового транзистора в равновесном состоянии.

¹ Наряду с основным ускоряющим полем, действующим на протяжении почти всей базы, имеется еще и тормозящее поле, сосредоточенное на небольшом участке вблизи эмиттерного перехода [78, 14]. Происхождение тормозящего поля связано с тем, что кривая $N_d(x)$ у реальных приборов не является монотонно спадающей, как показано на рис. 4-37, а имеет максимум при $x > 0$. В дальнейшем наличие тормозящего поля не учитывается, так как оно не меняет принципиальных особенностей дрейфовых транзисторов.

П О Л Н Ы Х концентраций можно записать в следующем виде:

$$\frac{d^2 p}{dx^2} - \frac{2\eta}{w} \frac{dp}{dx} - \frac{p}{L^2} = 0. \quad (4-122)$$

Безразмерный коэффициент η называется *коэффициентом неоднородности базы*¹:

$$\eta = \frac{w}{2L_d}. \quad (4-123a)$$

Для однородной базы, в которой $L_d = \infty$, получается $\eta = 0$. Если в формулу (4-120) подставить $x = w$ и положить $L_N = L_d$, то коэффициент неоднородности можно выразить через граничные концентрации примеси:

$$\eta = \frac{1}{2} \ln \frac{N_d(0)}{N_d(w)}. \quad (4-123б)$$

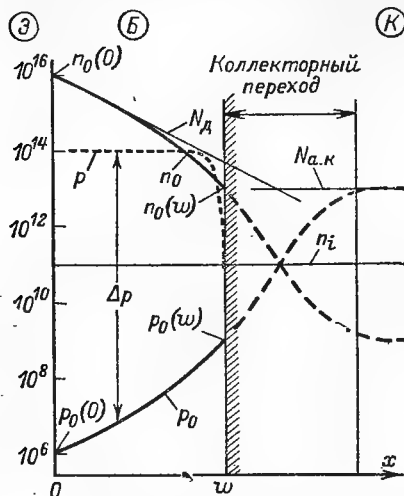


Рис. 4-39. Распределение равновесных концентраций носителей в базе дрейфового транзистора. Точками показаны распределение дырок при инжекции (см. рис. 4-40).

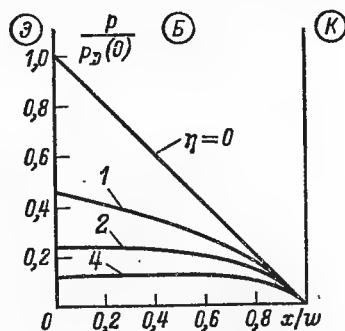


Рис. 4-40. Распределение концентрации инжектированных дырок в базе дрейфового транзистора при низких уровнях инжекции.

Отношение граничных концентраций $N_d(0)$ и $N_d(w)$ может быть весьма большим, но все же оно ограничено. Действительно, значение $N_d(0)$ должно удовлетворять условию высокого коэффициента инжекции $N_d(0) \ll N_{a,э}$ (где $N_{a,э}$ — концентрация акцепторных примесей в эмиттере), а значение $N_d(w)$, как видно из рис. 4-37, всегда превышает концентрацию $N_{a,к}$ в исходной пластине. Поскольку концентрация $N_{a,э}$ ограничена (сверху) предельной растворимостью примесей, а концентрация $N_{a,к}$ (снизу) — допустимым удельным сопротивлением коллекторного слоя (или достижимой степенью очистки полупроводника), то реальные отношения граничных концентраций $N_d(0)$ и $N_d(w)$ обычно лежат в пределах от 10^2 до 10^3 , а коэффициент неоднородности составляет $\eta = 2 \div 4$.

Общим решением уравнения (4-122) является функция

$$p(x) = A_1 e^{\xi_1 x} + A_2 e^{\xi_2 x}, \quad (4-124)$$

¹ Эту же величину часто называют *коэффициентом поля* (см. с. 85). Разница в определениях (4-123a) и (1-124) объясняется тем, что в § 1-13 толщина полупроводника считалась очень большой ($w \gg L$), тогда как в транзисторах толщина базы мала ($w \ll L$). Поэтому в первом случае параметр L_d было целесообразно сравнивать с величиной L , а во втором случае — с величиной w .

где корни характеристического уравнения имеют вид ¹:

$$\xi_{1,2} = \frac{1}{w} (\eta \pm \sqrt{\eta^2 + \zeta^2}). \quad (4-125)$$

В формуле (4-125) введено специальное обозначение для относительной толщины базы

$$\zeta = \frac{w}{L}.$$

У бездрейфовых транзисторов этот параметр определяет величину коэффициента переноса κ [см. (4-18)].

Теперь запишем граничные условия, с помощью которых можно найти коэффициенты A_1 и A_2 .

Первое граничное условие можно основать на том, что полный дырочный ток при $x = 0$ равен сумме диффузионной и дрейфовой составляющих. Используя выражения (1-72а) и (1-73а), учитывая соотношение (1-74) и подставляя значение E из (4-121), получаем:

$$-\left. \frac{dp}{dx} \right|_0 + \frac{p(0)}{L_n} = \frac{I_{\text{эп}}}{qDS}. \quad (4-126a)$$

В качестве второго граничного условия примем

$$p(w) = 0; \quad (4-126b)$$

оно соответствует значению $|U_K| \gg \Phi_T$ в общем выражении (4-166).

Используя корни (4-125) и найдя коэффициенты A_1 и A_2 с помощью граничных условий (4-126), можно представить стационарное распределение концентрации в следующем общем виде:

$$p(x) = \frac{I_{\text{эп}} w}{qDS} \frac{e^{\eta \frac{x}{w}} \operatorname{sh} \left[\sqrt{\eta^2 + \zeta^2} \left(1 - \frac{x}{w} \right) \right]}{\eta \operatorname{sh} \sqrt{\eta^2 + \zeta^2} + \sqrt{\eta^2 + \zeta^2} \operatorname{ch} \sqrt{\eta^2 + \zeta^2}}. \quad (4-127a)$$

В случае однородной базы ($\eta = 0$) выражение (4-127а), естественно, переходит в (4-176). При дополнительном условии $\zeta \ll 1$ получаем почти линейное распределение (4-17в). Этому распределению на рис. 4-40 соответствует кривая с параметром $\eta = 0$.

Если пренебречь рекомбинацией в базе, т. е. положить $\zeta = 0$, то распределение приобретает существенно более простой вид ²:

$$p(x) = \frac{I_{\text{эп}} w}{qDS} \frac{1 - e^{-2\eta \left(1 - \frac{x}{w} \right)}}{2\eta}. \quad (4-127b)$$

¹ В случае однородной базы, для которой $\eta = 0$, корни имеют значения $\xi_{1,2} = \pm 1/L$. При этом выражение (4-124) переходит в выражение (4-15) с точностью до члена p_0 , опущенного в уравнении (4-122).

² Формулу (4-127b) можно получить непосредственно из (4-127а) либо полагая $L = \infty$ в исходном уравнении (4-122), либо используя корни (4-125), которые при $\zeta = 0$ имеют значения $\xi_1 = 2\eta/w$ и $\xi_2 = 0$.

Первый множитель в выражениях (4-127) есть граничная концентрация неосновных носителей для бездрейфового транзистора с тонкой базой [см. (4-17в) при $x = 0$]. Эта концентрация выбрана в качестве основы масштаба на рис. 4-40 и обозначена через $p_D(0)$.

Из выражений (4-127) следует, что граничная концентрация у дрейфового транзистора всегда меньше, чем у бездрейфового: $p(0) < p_D(0)$.

Режим высокого уровня инжекции имеет у дрейфовых транзисторов свою специфику, которая состоит в следующем. Поскольку

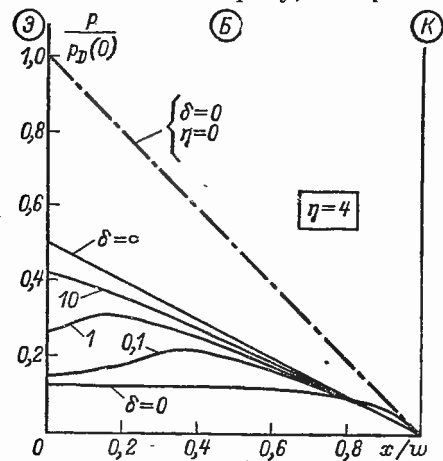


Рис. 4-41. Распределение неосновных носителей в базе дрейфового транзистора при высоких уровнях инжекции (штрихпунктиром показано распределение в бездрейфовом транзисторе при низком уровне инжекции).

равновесная концентрация основных носителей n_0 спадает вдоль оси x весьма резко (экспоненциально), а концентрация избыточных неосновных носителей Δp меняется сравнительно слабо (см. рис. 4-40), то уровень инжекции $\Delta p/n_0$ оказывается максимальным не на эмиттерной границе базы (как у бездрейфовых транзисторов), а вблизи коллектора (см. рис. 4-39). Отсюда следует, что с увеличением тока модуляция удельного сопротивления базы (см. § 2-8) происходит в направлении от коллектора к эмиттеру и что приэмиттерный слой базы модулируется слабее других. Поэтому коэффициент δ в выражении (2-69), характеризующий при-

эмиттерный уровень инжекции, у дрейфовых транзисторов редко превышает единицу даже при больших токах, тогда как у транзисторов с однородной базой значение δ достигает 10—100. Соответственно коэффициент инжекции [см. выражение (2-74)] у дрейфовых транзисторов зависит от тока сравнительно слабо.

Распределение инжектированных носителей при высоком уровне инжекции (рис. 4-41) [65] можно пояснить, исходя из следующих соображений. Как известно, модуляция удельного сопротивления базы связана с ростом концентрации основных носителей, в нашем случае — электронов. Вблизи коллектора, где модуляция максимальна (см. выше), концентрация электронов намного превышает равновесное значение. В результате распределение электронов на модулированном участке оказывается более равномерным, чем в равновесном состоянии, а собственное электрическое поле соответственно более слабым. Иначе говоря, модулированный участок в значительной мере лишен «дрейфовой специфики».

и подобен базе бездрейфового транзистора. В частности, на этом участке распределение дырок должно быть почти линейным¹.

Все сказанное подтверждается кривыми на рис. 4-41. Рассмотрим типичную кривую с параметром $\delta = 0,1$. Ее начальный участок почти горизонтален, поскольку здесь модуляция несущественна и дырки движутся в основном за счет дрейфа в собственном поле базы. Конечный участок кривой почти линеен, поскольку здесь модуляция максимальная, собственное поле практически отсутствует и дырки движутся под действием градиента концентрации с удвоенным коэффициентом диффузии [см. (2-72)]. Промежуточный участок характерен повышенной концентрацией дырок, поскольку в направлении от начального участка к конечному собственное поле убывает, а значит, уменьшается скорость носителей и происходит их накопление. Разумеется, повышенная концентрация p должна быть достаточной для того, чтобы дрейфовая составляющая тока могла компенсировать диффузионную составляющую, обусловленную положительным градиентом концентрации на этом участке.

Если дрейфовый транзистор работает в инверсном включении, то главная специфика состоит в том, что собственное поле E препятствует движению носителей, инжектированных через коллекторный переход. С аналитической точки зрения в уравнении (4-122) изменится знак второго члена, а в результирующих выражениях (4-127) коэффициент η следует считать отрицательным. Соответственно распределение $p(x)$ без учета рекомбинации (4-127б) принимает вид:

$$p(x) = \frac{I_{kp} w}{qDS} \frac{e^{2\eta \left(1 - \frac{x}{w}\right)} - 1}{2\eta}, \quad (4-128)$$

где координата x отсчитывается от коллектора к эмиттеру.

Кривые $p(x)$ для инверсного включения показаны на рис. 4-42. Как видим, они существенно отличаются от кривых $p(x)$ для нор-

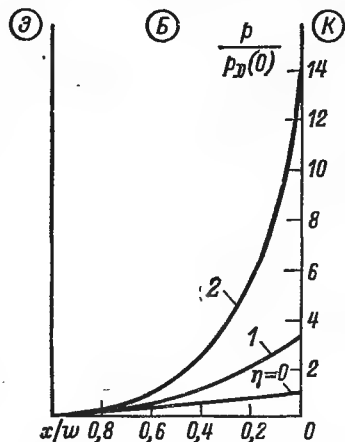


Рис. 4-42. Распределение концентрации инжектированных дырок в базе дрейфового транзистора при инверсном включении.

¹ Напомним, что в транзисторе с однородной базой диффузионный ток при высоком уровне инжекции составляет лишь половину полного тока [см. (2-72) и рис. 2-36, б]. Поэтому на рис. 4-41 асимптотическая прямая с параметром $\delta = \infty$ (общая для дрейфовых и бездрейфовых транзисторов) имеет вдвое меньший наклон, чем штрихпунктирная прямая с параметром $\delta = 0$, соответствующая транзистору с однородной базой. Общность асимптотической прямой $\delta = \infty$ для однородной и неоднородной баз обусловлена следующим: если концентрации избыточных носителей намного превышают концентрации равновесных, то характер распределения последних (а значит, и характер распределения примеси в базе) не оказывает существенного влияния на распределение избыточных носителей, и это распределение оказывается одинаковым в транзисторах с однородной и неоднородной базами.

мального включения. В частности, граничная концентрация $p(0)$ на инжектирующем контакте (коллекторе) оказывается значительно выше, чем при нормальном включении, для одного и того же тока.

Коэффициент переноса. В точке $x = w$, т. е. на границе коллекторного перехода, концентрация неосновных носителей равна нулю и, следовательно, дрейфовая составляющая тока отсутствует. Дифференцируя (4-127а) по x , подставляя производную dp/dx в выражение (1-73а), полагая $x = w$ и умножая на площадь S , получаем величину коллекторного тока. Поделив этот ток на I_{sp} находим коэффициент переноса в общем виде ¹:

$$\kappa = \frac{e\eta\sqrt{\eta^2 + \xi^2}}{\eta \operatorname{sh} \sqrt{\eta^2 + \xi^2} + \sqrt{\eta^2 + \xi^2} \operatorname{ch} \sqrt{\eta^2 + \xi^2}}. \quad (4-129a)$$

Упрощать это выражение, пренебрегая рекомбинацией в базе, нельзя, так как именно рекомбинация обуславливает отличие коэффициента κ от единицы. Поэтому, чтобы упростить выражение (4-129а), воспользуемся соотношением $\xi^2 \ll \eta^2$. Вынося из-под корня величину η , разлагая корень в ряд с точностью до второго члена и раскрывая функции sh и ch , можно после преобразований и упрощений записать коэффициент переноса в форме, аналогичной (4-196):

$$\kappa = \frac{1}{1 + \frac{1}{2} \left(\frac{w}{L} \right)^2 k(\eta)}. \quad (4-129б)$$

Поправочный множитель $k(\eta)$ имеет следующие точное и приближенное значения:

$$k(\eta) = \frac{2\eta - 1 + e^{-2\eta}}{2\eta^2} \approx \frac{1}{1 + \eta}.$$

Для типичных значений $\eta = 2 \div 4$ получается $k(\eta) = 0,33 \div 0,20$. При оценочных расчетах вполне приемлемо приближенное значение $k(\eta) = 0,25$.

Полагая $\gamma = 1$ и подставляя (4-129б) в (4-68), запишем выражение для коэффициента усиления в схеме ОЭ:

$$\beta = \frac{2L^2}{w^2} \frac{1}{k(\eta)} \approx \frac{2L^2}{w^2} (1 + \eta). \quad (4-130)$$

Сравнивая (4-129б) с (4-196) и (4-130) с (4-75б), приходим к выводу, что у дрейфовых транзисторов коэффициент переноса ближе к единице, чем у бездрейфовых, а коэффициент усиления при прочих равных условиях в несколько раз больше.

В случае инверсного включения в выражении (4-129б) следует умножить правую часть на отношение эффективных площадей эмиттера и коллектора, поскольку в реальных дрейфовых транзисторах

¹ Легко убедиться, что в случае $\eta = 0$ выражение (4-129а) переходит в выражение (4-18) для бездрейфового транзистора.

(см. рис. 4-36) значительная часть инжектированных коллектором носителей не попадает на эмиттер. Кроме того, в поправочном коэффициенте $k(\eta)$ (имеется в виду его точное значение) нужно изменить знак при η . Тогда

$$\alpha_I = \frac{S_n/S_k}{1 + \frac{1}{2} \left(\frac{w}{L} \right)^2 \left(\frac{e^{2\eta} - 2\eta - 1}{2\eta^2} \right)}. \quad (4-131)$$

Переходя к величинам α_I и β_I , следует иметь в виду, что коэффициент инжекции при инверсном включении дрейфовых транзисторов существенно меньше единицы ($\gamma_I \approx 1/2$) из-за сравнимых удельных сопротивлений коллектора и приколлекторного слоя базы. Поэтому, например, коэффициент β_I следует записывать в форме (4-75а):

$$\beta_I = \frac{\gamma_I (S_n/S_k)}{1 - \gamma_I (S_n/S_k) + \frac{1}{2} \left(\frac{w}{L} \right)^2 \left(\frac{e^{2\eta} - 2\eta - 1}{2\eta^2} \right)}. \quad (4-132)$$

Сравнивая это выражение с (4-130), убеждаемся, что $\beta_I \ll \beta$. Практически значение β_I лежит в пределах 0,1—1,0.

Динамические параметры. Движение носителей в слое базы при наличии поля характеризуется не только временем диффузии t_D , но и временем дрейфа $t_{др}$. Последнее определяется скоростью дрейфа μE и расстоянием w . Учитывая (4-121), а также соотношение (1-74), получаем:

$$t_{др} = \frac{w}{\mu E} = \frac{w}{D} L_d. \quad (4-133)$$

Сравнивая (4-133) и (2-88), приходим к выводу, что

$$\frac{t_{др}}{t_D} = \frac{2L_d}{w} = \frac{1}{\eta}. \quad (4-134)$$

Время дрейфа будет меньше, чем время диффузии, если $\eta > 1$ или, с учетом (4-123б), если

$$N_d(0)/N_d(w) > e^2 \approx 7.$$

Это условие всегда выполняется в дрейфовых транзисторах.

За счет дрейфа получается заметный выигрыш во времени прохождения носителей через базу в дополнение к тому выигрышу, который обусловлен малой толщиной базы. Этот дополнительный выигрыш можно оценить приближенно следующим образом.

Обозначим результирующее время пролета носителей через τ_α . За это время они пройдут путь $\sqrt{2D\tau_\alpha}$ за счет диффузии¹ и путь $\mu E\tau_\alpha$ за счет дрейфа. Сумма обеих величин есть, очевидно, толщина базы w :

$$\mu E\tau_\alpha + \sqrt{2D\tau_\alpha} = w.$$

¹ Такой вывод следует из структуры выражения (2-88), которое связывает расстояние w и время t_D .

Разделив обе части на ω и используя выражения (2-88), (4-133) и (4-134),

можно получить соотношение $\eta \frac{\tau_a}{t_D} + \sqrt{\frac{\tau_a}{t_D}} = 1$.

Отсюда постоянная времени τ_a выражается следующим образом:

$$\tau_a = t_D \frac{2\eta + 1 - \sqrt{4\eta + 1}}{2\eta^2}. \quad (4-135)$$

Строгий анализ должен исходить из нестационарного уравнения непрерывности (1-78а), которое можно решать операторным методом. Однако более простой и наглядный путь состоит в использовании связи между временем пролета носителей через базу, током эмиттера и зарядом в базе. В случае бездрейфовых транзисторов с тонкой базой эта связь выражается соотношением (4-62), которое позволяет определить время диффузии t_D как время «обновления» заряда в базе (см. с. 150). В случае дрейфовых транзисторов определим время пролета t_{np} из аналогичного соотношения:

$$t_{np} = \frac{\Delta Q}{I_{эп}}. \quad (4-136)$$

Неравновесный заряд ΔQ можно найти, интегрируя распределение (4-127б) в пределах от 0 до ω и умножая результат на элементарный заряд q и площадь S :

$$\Delta Q = \frac{I_{эп} \omega^2}{4D\eta^2} (2\eta - 1 + e^{-2\eta}). \quad (4-137)$$

Подставляя значение ΔQ в выражение (4-136) и выделяя время диффузии (2-88), получаем ¹:

$$t_{np} = t_D \frac{2\eta - 1 + e^{-2\eta}}{2\eta^2} \approx \frac{t_D}{1 + \eta}. \quad (4-138)$$

Как известно, суммарное время пролета t_D в бездрейфовых транзисторах складывается из времени $t_{зк}$ и постоянной времени фронта τ_k [см. (4-50)]. Следует ожидать, что у дрейфовых транзисторов время пролета t_{np} имеет такую же структуру. Однако без строгого анализа выделить составляющие $t_{за}$ и τ_a можно лишь с помощью некоторых упрощающих предположений. А именно, примем, что средняя скорость носителей в процессе диффузии и оной задержки есть $\omega/(0,2 t_D)$. Складывая эту величину с дрейфовой скоростью μE , деля толщину базы ω на сумму обеих скоростей и учитывая соотношения (4-133) и (4-134), получаем:

$$t_{за} = \frac{0,2 t_D}{1 + 0,2\eta}. \quad (4-139а)$$

Постоянная времени τ_a определится из соотношения $\tau_a = t_{np} - t_{за}$:

$$\tau_a = 0,8 t_D \frac{1 + 0,25\eta}{(1 + \eta)(1 + 0,2\eta)}. \quad (4-139б)$$

¹ Относительно аппроксимации, использованной в (4-138), см. с. 248.

Сравнивая формулы (4-139) с формулами (4-50), видим, что времена задержки у бездрейфовых и дрейфовых транзисторов различаются менее сильно, чем постоянные времени τ_α . Поэтому относительная задержка [фактор c , см. (4-60)] оказывается у дрейфовых транзисторов значительно больше:

$$c = \frac{t_{3\alpha}}{\tau_\alpha} = \frac{0,25(1+\eta)}{1+0,25\eta}. \quad (4-140)$$

На рис. 4-43 показаны зависимости временных параметров дрейфового транзистора от коэффициента неоднородности η . Из рисунка видно, что при значениях $\eta = 2 \div 3$ выигрыш в значении τ_α , обусловленный наличием поля, можно оценить примерно в 3—5 раз. Фактор задержки для дрейфовых транзисторов часто полагают равным 0,6, что соответствует значению $\eta \approx 3,5$.

Подводя итог, можно сказать, что дрейфовые транзисторы, несмотря на специфику в структуре и физических процессах, описываются той же системой параметров и теми же эквивалентными схемами, что и бездрейфовые. В частности, остаются в силе соотношения (4-84) для схемы ОЭ. Коррективы приходится вносить лишь в те формулы, где фигурирует время диффузии: его следует заменять на время пролета (4-138). К числу таких формул относятся выражения (4-63) для диффузионной емкости эмиттера и выражение (4-86) для предельной частоты. В случае дрейфовых транзисторов диффузионная емкость при прочих равных условиях существенно меньше:

$$C_{э,д} \approx \frac{t_{пр}}{r_э} \approx \frac{t_D}{r_э} \frac{1}{1+\eta}, \quad (4-141)$$

а предельная частота существенно больше:

$$\omega_T = \frac{\alpha}{t_{пр}} \approx \frac{\alpha}{t_D} (1+\eta). \quad (4-142)$$

Влияние барьерной емкости эмиттера (т. е. постоянной времени τ_v) учитывается с помощью тех же поправок, которые были подробно рассмотрены применительно к бездрейфовым транзисторам [см. (4-56) — (4-61)].

Постоянная времени τ_b в схеме с общим эмиттером остается равной времени жизни τ в силу тех же причин, о которых говорилось в связи с формулой (4-82). Однако значение времени жизни у дрейфо-

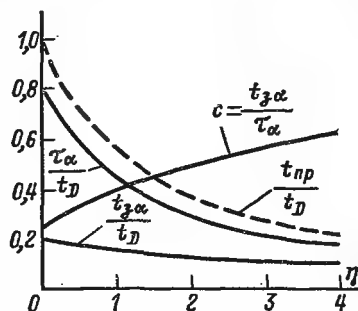


Рис. 4-43. Зависимости динамических параметров дрейфового транзистора от коэффициента неоднородности базы.

вых транзисторов значительно меньше (поскольку база сильнее легирована) и обычно составляет десятки наносекунд.

Предельные частоты современных дрейфовых транзисторов доходят до 5 ГГц (кремний) и 10 ГГц (германий), что соответствует сантиметровому диапазону длин волн. При этом отдаваемая мощность на частотах 1—2 ГГц достигает нескольких ватт. В импульсном режиме времена нарастания таких транзисторов не превышают 1 нс. Дальнейшее усовершенствование переходных и частотных свойств дрейфовых транзисторов должно сопровождаться одновременным усовершенствованием, а точнее, кардинальным изменением всех других элементов полупроводниковых схем наносекундного или СВЧ диапазонов. Наиболее адекватным решением этой проблемы является переход к микроэлектронным (интегральным) схемам.

4-13. ЭЛЕМЕНТЫ ТЕХНОЛОГИИ ТРАНЗИСТОРОВ

Полупроводниковая технология в теоретическом плане разработана сравнительно слабо и в этом отношении еще далека от уровня физики полупроводников и полупроводниковых приборов. Тем не менее целесообразно дать общее представление об основных этапах технологического процесса, поскольку такие сведения могут способствовать лучшему пониманию свойств, параметров и особенностей самих приборов. Более полное изложение технологии транзисторов можно найти в [74, 79—81].

Получение и очистка полупроводников. Качество полупроводниковых приборов в значительной степени зависит от качества исходных полупроводниковых материалов. Особую проблему при изготовлении полупроводников представляет их очистка. Так, в § 1-4 была отмечена возможность превращения полупроводника в полуметалл, если концентрация примеси достаточно велика. Обычно такое превращение происходит при относительном содержании примеси (0,01-0,001%), что характерно, например, для эмиттерных слоев диодов и транзисторов. Для сохранения характерных полупроводниковых свойств содержание примеси, как правило, должно лежать в пределах 0,0001%. Однако и эта исключительно малая цифра характерна лишь для полезной примеси. Содержание посторонних, а особенно вредных примесей должно быть еще на 2—3 порядка меньше¹.

Идеальным случаем была бы возможность получения абсолютно чистого (собственного) полупроводника, в который потом можно было бы добавлять необходимое количество полезной примеси. Практически получение «настоящего» собственного полупроводника невозможно, но методы современной металлургии позволяют получить исходные материалы с указанной выше степенью чистоты.

¹ Абсолютная величина концентрации примесей, разумеется, огромна. Например, 0,0000001% примесей означает концентрацию их больше, чем 10^{13} атомов в кубическом сантиметре.

Германий и кремний, полученные путем химического обработки их двуокисей, с точки зрения полупроводниковой техники считаются очень загрязненными, почти металлами. Методы специальной очистки их основаны на том, что растворимость большинства примесей гораздо больше в жидкой фазе, чем в твердой. Поэтому если постепенно охлаждать расплавленный германий или кремний, то в затвердевшей части будет меньше примесей, чем в остающейся жидкой части. Как говорят, примеси *оттесняются* в жидкую фазу. Отрезая от полностью затвердевшего слитка ту часть, которая затвердела последней (и в которой, следовательно, сконцентрирована основная масса примесей), и повторяя операцию, можно получить

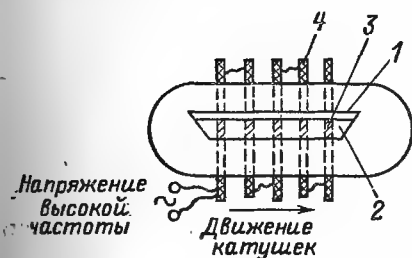


Рис. 4-44. Схема зонной плавки.

1 — тигель; 2 — твердая фаза полупроводника; 3 — жидкая фаза; 4 — подвижные катушки индукционного нагрева.

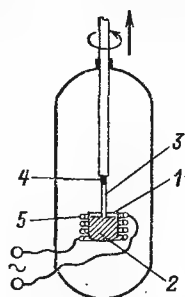


Рис. 4-45. Схема выращивания монокристалла (метод Чохральского).

1 — тигель; 2 — расплав полупроводника; 3 — вытягиваемый монокристалл; 4 — затравка; 5 — обмотка электропечи.

очень чистый материал. Реализацией описанного принципа являются такие распространенные методы очистки, как *зонная плавка* и *выращивание кристалла из расплава* (метод Чохральского) [79].

При зонной плавке (рис. 4-44) исходный слиток помещают в челночный тигель (из кварца или графита), относительно которого медленно перемещается группа из пяти-шести катушек, служащих для индукционного нагрева и расплавления тех участков слитка, которые находятся в данный момент под катушками. Одна «протяжка» слитка через группу катушек эквивалентна нескольким «протяжкам» через одну катушку. Иногда степень оттеснения примесей за одну «протяжку» оказывается все же недостаточной; тогда отрезав «примесную» часть слитка, операцию повторяют.

Метод выращивания кристалла (рис. 4-45) состоит в том, что в тигель с расплавленным полупроводником опускают (до соприкосновения с расплавом) монокристаллическую «затравку» из того же материала, а затем медленно поднимают ее вместе с постепенно нарастающим на ней новым монокристаллом. Для перемешивания расплава стержень с «затравкой» медленно вращают или заставляют слегка вибрировать. Метод выращивания обычно применяют после предварительной зонной плавки.

Все виды очистки полупроводников проводят в атмосфере инертных газов или водорода, которые также должны быть достаточно чистыми.

Контроль полупроводниковых материалов помимо металлографических испытаний включает ряд электрических измерений, в том числе измерения удельной проводимости (или сопротивления) и концентрации носителей [82].

Чаще всего измерение удельного сопротивления осуществляется методом четырех зондов [1, 4], а измерение концентраций — путем измерения э. д. с. Холла [15, 82].

Механическая и химическая обработка. Массивные кристаллические слитки полупроводников диаметром до 40—80 мм и длиной до 50—75 см разрезают на множество тонких пластин, на основе которых затем изготавливают отдельные приборы или интегральные схемы.

Резка слитка на пластины осуществляется либо с помощью комплекта тонких стальных полотен, либо с помощью «алмазных дисков». В первом случае собственно режущим агентом служат частички абразивного порошка, которые входят в состав суспензии, непрерывно поступающей к движущимся полотнам. Во втором случае режущим агентом является алмазная крошка, вкрапленная по периферии вращающегося металлического диска или, чаще, по внутренней кромке вращающегося кольца; охлаждение в этом случае осуществляется струей воды.

Оба указанных метода характерны значительными потерями полупроводникового материала. Поэтому осваиваются новые способы резки с помощью ультразвука, лазеров, электронного луча и т. п.

Во время резки слиток прочно закрепляют на неподвижном основании, причем очень важно обеспечить точное расположение слитка относительно полотен или дисков с тем, чтобы пластины имели необходимую кристаллографическую ориентацию. Контроль ориентации осуществляется рентгеновским методом. Как правило, пластины нарезаются в плоскости (111) или (100) (см. § 1-2).

Поверхность пластин после резки весьма неровная: размеры царапин, выступов и ямок иногда намного превышают размеры будущих структурных элементов. Поэтому перед началом основных технологических операций пластины многократно шлифуют, а затем полируют. Цель шлифовки помимо удаления механических дефектов состоит также в том, чтобы обеспечить необходимую толщину пластины (200—300 мкм), недостижимую при резке, и параллельность плоскостей.

Ш л и ф о в к а осуществляется на вращающихся шлифовальных кругах, на которых располагают сразу несколько пластин. Шлифующим агентом являются микропорошки, размер зерна которых выбирается все меньшим при каждом очередном цикле шлифовки, вплоть до 3—5 мкм. В большинстве случаев шлифовка

бывает односторонней, т. е. каждая из двух плоскостей пластины шлифуется раздельно.

По окончании шлифовки на поверхности все же остается механически нарушенный слой толщиной от нескольких микрон и выше, под которым расположен еще более тонкий, так называемый «физически нарушенный» слой. Последний характерен наличием «незримых» искажений кристаллической решетки и механических напряжений, возникающих в процессе шлифовки.

Полировка состоит в удалении обоих нарушенных слоев и снижении неровностей поверхности до уровня, свойственного оптическим системам и составляющего сотни, а иногда десятки ангстрем (тысячные доли микрометра.) Чаще всего используется химическая полировка (травление), т. е. по существу растворение поверхностного слоя полупроводника в тех или иных реактивах. Обязательными компонентами таких реактивов являются окислитель (обычно азотная кислота) и растворитель образующегося окисла (обычно плавиковая кислота). Кроме этих компонентов в состав травителей входят ускорители и замедлители реакции. Выступы и трещины на поверхности стравливаются быстрее, чем основной материал, и в целом поверхность выравнивается.

Иногда химическую полировку сочетают с предварительной механической, осуществляемой с помощью микропорошков (в виде паст или суспензий) с размером зерна 0,1—0,3 мкм.

Достигаемая в процессе шлифовки и полировки параллельность плоскостей пластины составляет единицы и даже доли микрона на сантиметр длины.

Помимо травления важнейшим химическим процессом в полупроводниковой технологии является очистка поверхности от загрязнений органическими веществами, особенно жирами. Такую очистку приходится проводить неоднократно, в частности перед каждым травлением, поскольку скорость травления резко уменьшается в местах загрязнений.

Очистку и обезжиривание проводят в органических растворителях (толуол, ацетон, этиловый спирт и др.) при повышенной температуре. В последнее время ванну с растворителем часто укрепляют на магнитострикторе, обеспечивающем ультразвуковые колебания в растворе и тем самым интенсивное его перемешивание.

Травление, очистка и многие другие процессы сопровождаются отмывкой пластин в деионизованной воде. Деионизация осуществляется в специальных установках путем пропускания предварительно дистиллированной воды через пористые смолы, в которых происходит адсорбция растворенных ионов. Степень деионизации оценивается по удельному сопротивлению воды, которое должно лежать в пределах 10—20 МОм·см и выше (удельное сопротивление дистиллированной воды не превышает 1—2 МОм·см).

Эпитаксия. Описанные выше технологические процессы являются подготовительными и вспомогательными в том смысле, что сами по себе они не обеспечивают создания тех неоднородностей

в кристалле (например, *p-n* переходов), которые необходимы для работы диодов и транзисторов.

Эпитаксию можно определить как процесс наращивания монокристаллического слоя на подложку при сохранении кристаллографической ориентации последней.

Эпитаксия занимает промежуточное место между основными и вспомогательными процессами. В принципе она может обеспечить создание л о к а л ь н ы х неоднородностей, однако это направление еще слабо разработано и в настоящее время эпитаксия обычно используется как способ получения пластин — «заготовок» в виде тонких слоев однородного полупроводника на более толстой подложке.

Процесс эпитаксии применительно к кремнию состоит в следующем. Монокристаллическая кремниевая пластина помещается в поток водорода, содержащий пары галоидного соединения кремния (обычно SiCl_4). При высокой температуре (1200°C и выше) на поверхности пластины происходит реакция

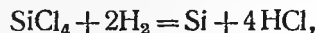


Рис. 4-46. Примеры эпитаксиальных слоев на разных типах подложки.

В результате которой на подложке осаждается чистый кремний, а пары HCl уносятся потоком водорода. Эпитаксиальный слой осажденного кремния монокристаллический и имеет ту же кристаллографическую ориентацию, что и подложка.

Если к парам тетрахлорида кремния добавить пары галоидных соединений бора (BBr_3) или фосфора (PCl_3), то эпитаксиальный слой будет иметь уже не собственную, а соответствующую дырочную или электронную проводимость, поскольку в ходе химической реакции в кремний будут внедряться акцепторные атомы бора или донорные атомы фосфора.

Таким образом, эпитаксия позволяет выращивать монокристаллические слои любого типа проводимости и любого удельного сопротивления на подложке, обладающей тоже произвольными параметрами как по типу, так и по величине проводимости (рис. 4-46).

Граница между эпитаксиальным слоем и подложкой не получается идеально резкой, так как примеси в процессе эпитаксии частично диффундируют из одного слоя в другой. Поэтому попытки создать многослойные эпитаксиальные структуры лишь в последние годы увенчались успехом. Основную роль в настоящее время играет однослойная эпитаксия, при помощи которой возможно получение тонких однородных слоев до одного микрона, что невозможно осуществить иными средствами¹.

¹ Даже если бы шлифовальные и полировальные установки позволяли достигнуть такой толщины, полученные пластины не допускали бы никакой последующей обработки из-за механической хрупкости.

Диффузия. Внесение примесей в исходную пластину (или в эпитаксиальный слой) путем диффузии при высокой температуре является в настоящее время основным способом создания диодных и транзисторных структур.

Диффузия может быть локальной и общей. В первом случае она осуществляется на определенных участках пластины через специальные маски (рис. 4-47, а, б), во втором — по всей поверхности (рис. 4-47, в) *.

Диффузию можно проводить и однократно, и многократно (двойная, тройная диффузия). Например, в исходную пластину n -типа можно во время 1-й диффузии внедрить акцепторную примесь и получить p -слой, а затем, во время 2-й диффузии, внедрить в полученный слой (на меньшую глубину) донорную примесь и тем самым обеспечить трехслойную структуру (рис. 4-47, г).

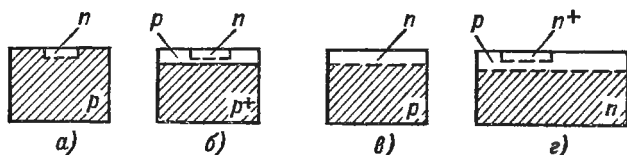


Рис. 4-47. Примеры использования процесса диффузии примесей.

а — локальная диффузия в пластину; б — локальная диффузия в эпитаксиальный слой; в — общая диффузия на одной из поверхностей пластины; г — двойная диффузия — общая (p -слой) и локальная (n^+ -слой).

Источником примеси — диффузанта может быть либо жидкость, либо газ (пар). В первом случае поверхность пластины контактирует с расплавом, содержащим в качестве компонента необходимую примесь, во втором случае — с парами примеси или с потоком инертного газа-носителя, содержащего пары примеси. Второй метод имеет большее распространение.

На рис. 4-48 иллюстрируется так называемый *метод открытой трубы* для получения диффузионных слоев. Вдоль трубы с небольшой скоростью непрерывно проходит поток нейтрального газа-носителя, например аргона. В 1-й высокотемпературной зоне к этому потоку примешиваются пары диффузанта, полученные сублимацией из твердого источника. Попадая во 2-ю высокотемпературную зону, где расположены пластины полупроводника, молекулы диффузанта (например, PCl_3) диссоциируют. Атомы полезной примеси (в нашем примере — фосфор) адсорбируются и диффундируют в пластины на определенную глубину, а другие составляющие диффузанта (в нашем примере — хлор) уносятся

* Общая диффузия приводит к образованию тонкого диффузионного слоя, который отличается от эпитаксиального неоднородным (по глубине) распределением примеси (см. рис. 4-37, кривая N_d).

газм-носителем из зоны. Нагрев источника диффузанта и пластины осуществляется локальным высокочастотным полем с помощью внешних катушек.

Диффузия примесей имеет под собой ту же теоретическую базу, что и диффузия подвижных носителей заряда. Существенное отличие состоит, конечно, в отсутствии рекомбинации, а с количественной стороны — в несравненно меньших коэффициентах диффузии, а значит, и скоростях движения¹.

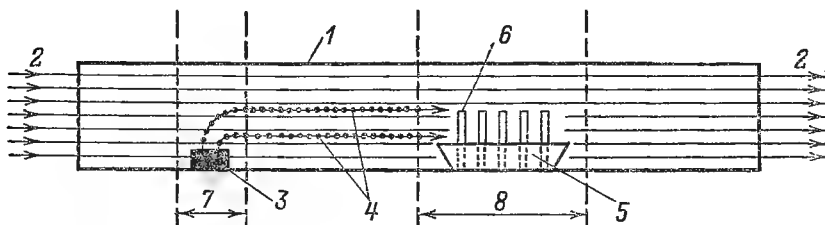


Рис. 4-48. Схема двухзонной диффузионной печи.

1 — кварцевая труба; 2 — поток газа-носителя; 3 — источник диффузанта; 4 — пар диффузанта; 5 — лодочка (челюстной тигель); 6 — пластина полупроводника, поставленная на ребро; 7 — 1-я высокотемпературная зона; 8 — 2-я высокотемпературная зона.

В отсутствие рекомбинации ($\tau = \infty$) любое из уравнений диффузии (1-79) применительно к концентрации примеси N запишется в виде 2-го закона Фика:

$$\frac{\partial^2 N}{\partial x^2} = \frac{1}{D} \frac{\partial N}{\partial t}, \quad (4-143a)$$

где $N = N(x; t)$ — распределение концентрации примеси.

В начальный момент времени можно считать концентрацию вводимой примеси в пластине равной нулю; тогда изображением производной $\partial N / \partial t$ будет просто sN , а уравнение (4-143a) в операторной форме будет следующим:

$$\frac{\partial^2 N}{\partial x^2} - \frac{s}{D} N = 0. \quad (4-143b)$$

Общим решением этого уравнения является функция

$$N(x; s) = A_1 e^{-x\sqrt{s/D}} + A_2 e^{x\sqrt{s/D}}. \quad (4-144)$$

Если глубина диффузии невелика по сравнению с толщиной пластины (или эпитаксиального слоя), то в качестве 1-го граничного условия можно принять:

$$N(\infty, s) = 0. \quad (4-145)$$

¹ Например, в кремнии коэффициенты диффузии бора и фосфора составляют при $T = 1250^\circ\text{C}$ около 10^{-11} см²/с. С изменением температуры коэффициенты диффузии резко меняются (для указанных примесей они удваиваются при $\Delta T = 30^\circ\text{C}$). Это обстоятельство заставляет контролировать температуру в диффузионных печах с точностью до $\pm 1^\circ\text{C}$ и менее.

Тогда $A_2 = 0$ и вместо (4-144) получаем:

$$N(x, s) = A_1 e^{-x\sqrt{s/D}}. \quad (4-146)$$

Второе граничное условие может иметь два варианта в зависимости от типа источника примеси. Различают неограниченный и ограниченный источники. Рассмотрим оба варианта поочередно.

В случае неограниченного источника примеси в качестве 2-го граничного условия принимают:

$$N(0; s) = N_0 = \text{const}. \quad (4-147)$$

Тогда $A_1 = N_0$ и, следовательно,

$$N(x; s) = N_0 e^{-x\sqrt{s/D}}. \quad (4-148)$$

Оригиналом такого изображения является дополнительная функция ошибок (см. сноску на с. 84)

$$N(x, t) = N_0 \operatorname{erfc}\left(\frac{x}{2\sqrt{Dt}}\right). \quad (4-149)$$

Эта зависимость проиллюстрирована на рис. 4-49, а. Дифференцируя (4-149) по x , получаем градиент концентрации:

$$N' = \frac{dN}{dx} = -\frac{N_0}{\sqrt{\pi Dt}} e^{-x^2/4Dt}. \quad (4-150)$$

В начальной точке $x = 0$ этот градиент выражается формулой

$$N'(0) = -\frac{N_0}{\sqrt{\pi Dt}}. \quad (4-151)$$

Поделив N_0 на $N'(0)$, получим некоторую условную глубину диффузии L_N (рис. 4-49, а), которая является аналогом длины L_N , свойственной экспоненциальному распределению (4-120). В данном случае

$$L_N = \sqrt{\pi} \sqrt{Dt} \approx 1,77 \sqrt{Dt}. \quad (4-152)$$

Как видим, глубина диффузии примеси находится в прямой зависимости от коэффициента диффузии и времени, в течение которого ведется процесс.

В случае ограниченного источника примеси в качестве 2-го граничного условия принимают:

$$\int_0^{\infty} N(x) dx = \bar{N} = \text{const}, \quad (4-153)$$

где $\bar{N}$ — полное количество атомов примеси (на единицу

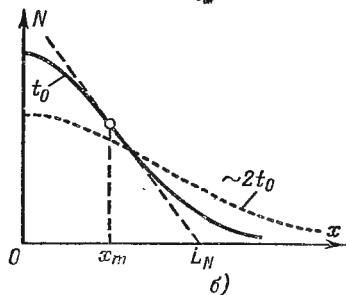
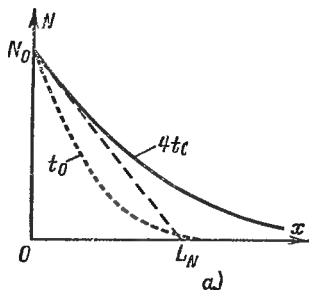


Рис. 4-49. Распределение примеси в пластине для двух моментов времени при диффузии из неограниченного (а) и ограниченного (б) источников.

площади), которое остается постоянным в процессе диффузии¹. Подставляя (4-146) в (4-153) и выполняя интегрирование, получаем:

$$A_1 = \bar{N} \sqrt{s/D},$$

тогда изображение (4-146) принимает вид:

$$N(x; s) = \bar{N} \sqrt{\frac{s}{D}} e^{-x \sqrt{s/D}}. \quad (4-154)$$

Оригиналом этого изображения будет *функция Гаусса*:

$$N(x; t) = \frac{\bar{N}}{\sqrt{\pi Dt}} e^{-x^2/4Dt}, \quad (4-155)$$

которая проиллюстрирована на рис. 4-49,б.

Начальная производная функции (4-155) равна нулю. Следовательно, глубину диффузии нельзя определить по аналогии с (4-152). Поэтому определим глубину диффузии иначе, в более общем виде. А именно, найдем координату x_m , при которой производная $N'(x_m)$ максимальна, и будем считать глубиной диффузии сумму отрезков x_m и $N(x_m)/|N'(x_m)|$ (см. рис. 4-49,б). При таком определении

$$L_N = 2\sqrt{2} \sqrt{Dt} \approx 2,82 \sqrt{Dt}. \quad (4-156)$$

Форма полученного выражения совпадает с формой (4-152), а различие оказывается чисто количественным и не очень существенным.

Скорость распространения примеси в процессе диффузии можно оценить, дифференцируя (4-152) или (4-156) по t . В обоих случаях получается примерно одинаковый результат:

$$v_N = \frac{dL_N}{dt} \approx \sqrt{\frac{D}{t}}. \quad (4-157)$$

Как видим, скорость распространения примеси убывает с течением времени, т. е. диффузия примеси идет наиболее интенсивно в начальном интервале.

При заданной (желательной) глубине диффузии соответствующее время процесса оценивается из формул (4-152) и (4-156):

$$t_{\text{диф}} \approx \frac{L_N^2}{4D}. \quad (4-158)$$

¹ Практически величина $\bar{N}$ обеспечивается путем предварительной диффузии («загонки») примеси на небольшую глубину из неограниченного внешнего источника, после чего источник отключается и следует автономная «разгонка» накопленных атомов. Поэтому величину $\bar{N}$ можно оценить с помощью распределения (4-149). Она оказывается равной $(2/\sqrt{\pi}) N_0 \sqrt{Dt_1}$, где t_1 — время «загонки».

Например, если $L_N = 2$ мкм и $D = 10^{-11}$ см²/с, то $t_{\text{диф}} = 1000$ с. Такое время считается приемлемым¹.

Основные технологические циклы. Технология полупроводниковых приборов быстро прогрессирует. Еще 15—17 лет назад господствовала *сплавная* технология. Затем появились *диффузионно-сплавная* и близкая к ней *мезатеchnология*. В конце 50-х годов была предложена *планарная* технология, которая в сочетании с *эпитаксиальной* технологией, разработанной около 15 лет назад, стала и остается основой производства кремниевых приборов и интегральных схем. Таким образом, в настоящее время имеется несколько различных технологических циклов, сочетающих те или иные процессы и методы.

Следует подчеркнуть, что современные диоды и транзисторы весьма редко изготавливаются в индивидуальном порядке, как это было в эпоху сплавной технологии. Основой современного производства полупроводниковых приборов является *групповой метод* изготовления, при котором на пластине в едином технологическом цикле создается сразу большое количество (200—300 и более) однотипных структур, а затем уже пластина разрезается на отдельные кристаллы и осуществляются «посадка на ножку», присоединение выводов и заключение каждого кристалла в корпус. Групповой метод дает не только экономические преимущества, но обеспечивает и улучшение технических параметров приборов, поскольку удается существенно уменьшить разброс параметров, размеры структур и т. д.

Одним из важных следствий внедрения группового метода явилось развитие *техники масок*². Назначение маски — защитить основную поверхность пластины и обеспечить проникновение примеси, травителя и других веществ только в определенные участки пластины, где расположены будущие структуры или их элементы. Для этого в маске делаются отверстия («окна») той или иной формы в количестве, соответствующем намеченному числу структур. В случае кремния роль маски с успехом играет слой его двуокиси, специально выращиваемый на поверхности (термическое окисление в атмосфере кислорода или паров воды). Окна в слое SiO₂ осуществляются методом *фотолитографии* (см. ниже). Окисный слой германия получить трудно, к тому же он оказывается весьма нестойким и непрочным. Поэтому в случае германия приходится специально наносить на поверхность защитный слой другого вещества (обычно того же SiO₂).

Одиночные маски используются редко. Чаще приходится иметь дело с комплектом масок, каждая из которых имеет свой

¹ Обычно выбирают такую температуру диффузии, которая обеспечивает время процесса не менее 10—20 мин. В этом случае можно получить заданную толщину диффузионного слоя с высокой точностью, поскольку прекращение нагрева с погрешностью, составляющей несколько, даже десятки секунд, оказывается малосущественным.

² В зависимости от назначения и способа изготовления маски называют *трафаретами, шаблонами, фотошаблонами*.

рисунок отверстий. Например, через 1-ю маску осуществляется диффузия базовых слоев p -типа, через 2-ю маску — диффузия эмиттерных слоев n -типа и т. п. Естественно, что отверстия в разных масках должны быть весьма точно согласованы, а при наложении очередной маски должно быть обеспечено *совмещение* ее отверстий с рисунком, полученным в результате предыдущего процесса.

Сплавная технология. Рассмотрим сплавной метод на примере германиевого диода, у которого эмиттером является слой p -типа (рис. 4-50). В этом случае за основу берется пластинка германия n -типа. На нее накладывают таблетку из акцепторного материала, обычно индия. После этого пластинку с таблеткой помещают в вакуумную или водородную печь и нагревают до такой температуры, при которой таблетка и прилегающий к ней слой пластинки расплавляются и образуют сплав определенного состава. Для комбинации германий — индий температура

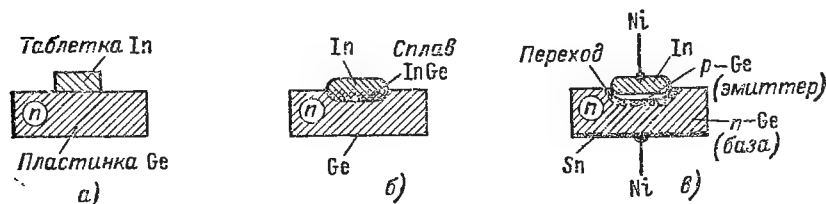


Рис. 4-50. Этапы изготовления сплавного диода.

а — исходные компоненты; б — процесс сплавления; в — результирующая структура диода.

нагрева составляет $450\text{—}550^\circ\text{C}$. Через несколько минут выдержки нагрев выключают, и сплав начинает застывать. При этом на дне капли образуется тонкий рекристаллизованный слой германия p -типа, а остальная часть застывшей капли состоит из почти чистого индия, который образует с p -слоем омический контакт. К этому контакту припаивают внешний вывод, обычно в виде никелевой проволоочки. На нижнюю грань пластинки наносят слой олова (которое дает омический контакт с германием n -типа) и к этому слою припаивают второй внешний вывод.

Четкая граница, образующаяся между жидким раствором InGe и исходной пластинкой в период нагрева¹, сохраняется после затвердевания. Поэтому сплавной метод обеспечивает получение ступенчатых p - n переходов. Удельное сопротивление рекристаллизованного слоя оказывается очень малым (примерно $0,001\text{—}0,01\text{ Ом} \cdot \text{см}$), и он играет роль эмиттера по отношению к гораздо более высокоомной исходной пластинке — базе.

При изготовлении сплавных транзисторов таблетку акцепторного материала накладывают на обе поверхности n -пластинки (с помощью специальной кассеты). Размер коллекторной таблетки несколько больше, чем эмиттерной. В результате после термической обработки получается структура, показанная на рис. 4-3.

Мезотехнология. Название этого технологического метода связано со специфической формой диодных или транзисторных структур, которая получается на одном из последних этапов групповой обработки пластины перед ее разрезанием на отдельные приборы.

¹ Четкость границы обусловлена сравнительно низкой температурой расплава InGe, при которой диффузия индия в смежный слой n -германия практически не происходит (из-за малого значения D , см. сноску на с. 258).

А именно, после избирательного травления структуры имеют форму конических выступов — «столиков», которые и называют «мезами» (рис. 4-51). Рассмотрим цикл мезатехнологии на примере германиевого транзистора.

За основу берут пластину p -типа и путем диффузии из газовой среды получают общий базовый слой n -типа толщиной 2—3 мкм (см. рис. 4-47, в). На одной из поверхностей пластины диффузионный слой стравливают, с тем чтобы в дальнейшем осуществить здесь омический контакт коллектора, а оставшийся диффузионный слой подвергают дальнейшей обработке, которая состоит в следующем.

Пластину накрывают «маской» — тонкой биметаллической пластинкой, в которой имеются прорезы на тех местах, где в будущем

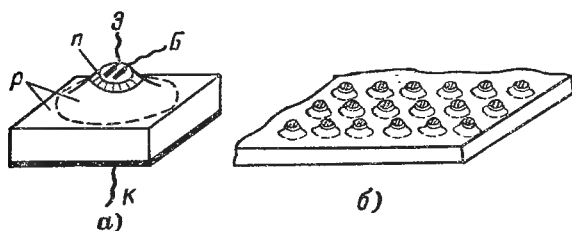


Рис. 4-51. Структура мезатранзистора.

а — отдельный транзистор; б — партия мезатранзисторов до разрезания пластины.

должны получиться эмиттеры (эти прорезы имеют размеры 25×100 мкм и число их достигает 100—200 на 1 см^2 площади). Пластины полупроводника, покрытую маской, помещают в вакуумную камеру и распыляют в ней (путем нагрева) алюминий. Сквозь прорезы маски алюминий осаждается на базовом n -слое, образуя узкие полоски.

Смещая ту же маску на 10—20 мкм и распыляя сплав сурьмы с золотом, получают полоски сплава рядом с полосками алюминия.

Пластину с нанесенными полосками нагревают, с тем чтобы материал полосок сплавился с n -слоем; в результате под алюминиевыми полосками получают тонкие (десятые доли микрона) слои эмиттеров p -типа, а под сурьмяно-золотыми полосками — омические контакты баз.

Накладывают новую маску с круглыми или прямоугольными отверстиями (размером 150—200 мкм). Эту маску располагают так, чтобы сквозь отверстия были видны ранее нанесенные полоски, и путем распыления покрывают область полосок защитным слоем¹.

¹ В последнее время вместо накладных масок используют метод фотолитографии (см. ниже).

Пластины с пятнышками защитного слоя помещают в травитель, который вытравливает n -слой и частично коллекторный p -слой в промежутках между пятнышками.

С полученной мезаструктуры (рис. 4-51, б) удаляют защитные пятнышки и разрезают ее на отдельные элементы.

К эмиттерной и базовой полоскам путем термокомпрессии (см. ниже) прикрепляют тонкие усики-выводы, а коллекторный слой припаивают к молибденовой пластинке с примесью золота. Осуществление выводов эмиттера и базы является самой сложной и тонкой операцией всего цикла.

На рис. 4-51, а видно что мезатранзистор характерен малыми площадями эмиттерного и коллекторного переходов и одновременно большой площадью коллекторного слоя. Эти особенности обеспечивают малое значение барьерных емкостей и сравнительно малое сопротивление коллекторного слоя, хотя большая толщина последнего является все же ограничивающим фактором и в ряде случаев не позволяет довести значение $r_{кк}$ до желательного минимума.

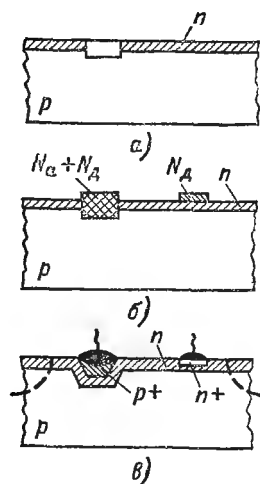


Рис. 4-52. Этапы изготовления диффузионно-сплавного транзистора.

а — пластина с лункой;
б — пластина с нанесенными сплавами; в — структура после термической обработки (пунктиром показаны контуры будущего мезавыступа).

Вариантом мезатехнологии является изготовление диффузионно-сплавных транзисторов. В этом варианте после реализации диффузионного базового n -слоя протравливают достаточно глубокие лунки сквозь весь базовый слой, несколько углубляясь в исходную p -пластину (рис. 4-52, а). Травление лунок (по числу будущих транзисторов) осуществляется методом фотолитографии (см. ниже). В полученные лунки тоже через маску, электрохимическим способом вводят комплексный сплав, содержащий одновременно две примеси — акцепторную и донорную, а рядом с лунками, непосредственно на диффузионный n -слой, наносят полоски донорного сплава (рис. 4-52, б).

Затем пластину помещают в печь и нагревают до температуры, близкой к температуре плавления германия (около 900°C). При такой температуре сплавы не только переходят в жидкое состояние (как при сплавной технологии), но и имеет место диффузия примесей из жидкой фазы в прилегающую твердую фазу. При этом комплексный характер сплава, находящегося в лунках, обеспечивает одновременное образование двух слоев — базового и эмиттерного благодаря резко различным коэффициентам диффузии донорной и акцепторной примесей в германии: донорная примесь «обгоняет» акцепторную. Что касается сплава, помещенного рядом с лунками, то диффузия доноров из него обеспечивает образование n^+ -слоя и тем самым омический контакт с исходным диффузионным n -слоем. Последний, как видно из рис. 4-52, в, играет роль соединительного слоя между базовым контактом и активной областью базы.

По завершении описанных процессов осуществляется избирательное травление поверхности пластинки и тем самым получается мезарельеф, подобный тому, который показан на рис. 4-51, б.

Сравнивая оба описанные варианта, можно заметить, что в первом случае базовый и эмиттерный слои получают на разных этапах технологического цикла (диффузия базы, затем сплавление эмиттера), а во втором случае на одном и том же этапе (диффузия базы плюс сплавление эмиттера).

Планарная технология. С конструктивной точки зрения планарные приборы характерны тем, что все их рабочие слои выходят на одну и ту же поверхность пластины и соответственно все электроды расположены в одной плоскости (см. рис. 4-36, б), тогда как мезаприборы имеют выводы в двух плоскостях (рис. 4-36, а).

Хотя указанная особенность непосредственно соответствует термину «планарный», обычно планарную технологию понимают более узко, как технологический цикл создания кремниевых приборов и интегральных схем методами локальной диффузии с использованием оксидных масок.

Маски в виде двуокиси SiO_2 , получаемой путем термического окисления поверхности кремниевой пластины, обладают следующими преимуществами.

1. Оксидная маска органически связана с поверхностью пластины, обеспечивая прочный контакт с нею и исключая проникновение диффузанта в зазор между маской и поверхностью.

2. Толщина оксидной маски (около микрона) достаточна для надежной защиты соответствующих участков пластины от проникновения диффузанта.

3. Оксидный слой одновременно с функцией маскирования выполняет функцию защиты поверхности (а значит, и p - n переходов, выходящих на поверхность) от влияния различных внешних факторов. В случае сплавной и мезатеchnологии для этого приходится использовать специальные средства — защитные покрытия.

Оксидные маски изготовляют методом фотолитографии, который имеет более общее значение в полупроводниковой технологии и описывается в следующем разделе.

Цикл изготовления планарного n - p - n транзистора иллюстрируется на рис. 4-53. За основу берется пластина кремния n -типа, которая в результирующей структуре играет роль коллектора. В эту пластину через 1-ю оксидную маску осуществляется диффузия акцепторной примеси (обычно бора) и получается слой p -базы. Затем через 2-ю оксидную маску осуществляется диффузия донорной примеси (обычно фосфора) и получается эмиттерный слой. Наконец с помощью 3-й оксидной маски осуществляют алюминиевые омические контакты ко всем трем слоям и далее присоединяют к этим контактам тонкие проволочки в качестве выводов¹.

В описанном варианте исходную пластину выбирают с достаточно высоким удельным сопротивлением, чтобы обеспечить необходимое пробивное напряжение коллекторного перехода. При

¹ Обычно такие проволочки, тоже алюминиевые (диаметром около 30 мкм), присоединяют методом *термокомпрессии*, т. е. путем сильного прижатия при высокой температуре (200—300 °C) на специальных установках.

этом сопротивление $r_{\text{кк}}$ коллекторного слоя в целом оказывается чрезмерно большим. Выходом из положения является планарно-эпитаксиальная технология, т. е. использование высокоомного эпитаксиального слоя на низкоомной подложке (рис. 4-54, а) или — в технологии интегральных схем — использование «скрытого» n^+ -слоя (рис. 4-54, б), получаемого на дне лунки до изготовления основной структуры. В обоих случаях ток протекает в горизонтальном направлении по низкоомному слою, тогда как коллекторный p - n переход расположен в высокоомном слое (т. е. пробивное напряжение достаточно велико).

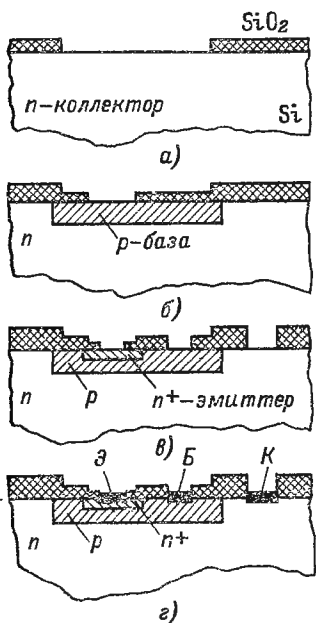


Рис. 4-53. Этапы изготовления планарного транзистора.

а — пластина перед диффузией базы; б — пластина перед диффузией эмиттера; в — пластина перед нанесением контактов; г — результирующая структура транзистора.

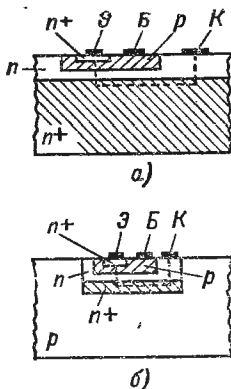


Рис. 4-54. Способы уменьшения поперечного сопротивления коллекторного слоя. Точками показан путь тока.

а — низкоомная подложка с высокоомным эпитаксиальным слоем; б — «скрытый» низкоомный n^+ -слой.

Фотолитография. Процесс фотолитографии [83] неоднократно упоминался как одно из важных средств современной полупроводниковой техники. Сущность фотолитографии применительно к кремниевой планарной технологии состоит в следующем (рис. 4-55).

На поверхность двуокиси SiO₂ наносится равномерный слой фотоэмульсии, так называемого *фоторезиста*. Сверху на слой фоторезиста накладывается стеклянная маска с прозрачными и зачерненными участками — *фотошаблон*. Сквозь эту маску засвечивают фоторезист ультрафиолетовым светом, как при обычной контактной фотопечати. После этого пластину с фоторезистом проявляют; в процессе проявления засвеченные участки фоторезиста стравливаются, и в этих местах обнажается поверхность двуокиси кремния.

Оставшийся (незасвеченный) слой фоторезиста подвергают термическому дублированию — полимеризации, в результате чего этот

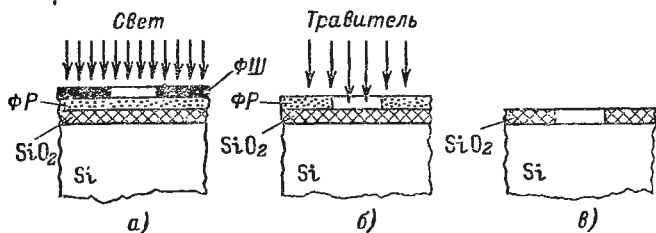


Рис. 4-55. Основные этапы фотолитографии по кремнию.

а — процесс засветки фоторезиста (ФР) через фотошаблон (ФС); *б* — травление оксидной пленки через «окно» в проявленном и задублированном фоторезисте; *в* — пластина после удаления фоторезиста.

слой становится нечувствительным к химическим травителям. Поэтому, когда на следующем этапе пластину подвергают травлению, растворяются лишь обнаженные участки двуокиси кремния, вплоть до поверхности самой пластины, вследствие чего в оксидной маске получается необходимая совокупность «окон», через которые в дальнейшем проводят локальную диффузию.

Заключительным этапом фотолитографии является удаление задублированного слоя фоторезиста, после чего пластина с оксидной маской готова к использованию для проведения диффузии.

Изготовление накладных фотошаблонов представляет самостоятельную, достаточно сложную задачу, поскольку размеры «окон» в настоящее время нередко составляют всего несколько микрон, а следовательно, края этих «окон» должны выполняться с еще большей точностью. Цикл производства фотошаблонов (рис. 4-56) начинается с вычерчивания оригинала в масштабе 200 : 1 и более по отношению к реальным размерам будущего рисунка. Такое вычерчивание (или чаще, выцарапывание на тонком непрозрачном слое) осуществляется на специальных прецизионных установках — координатографах. Далее следует промежуточный отъем — фотографирование оригинала с уменьшением размеров в 20—40 раз. Полученный негатив использу-

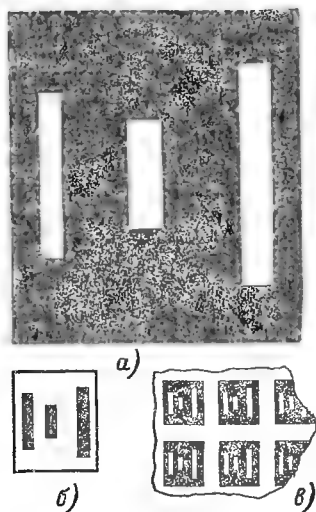


Рис. 4-56. Этапы изготовления фотошаблона под «окна» для омических контактов (рис. 4-53, в). Масштабы не выдержаны.

а — оригинал; *б* — промежуточный негатив; *в* — участок фотошаблона после мультипликации.

Далее следует промежуточный отъем — фотографирование оригинала с уменьшением размеров в 20—40 раз. Полученный негатив использу-

ется для размножения изображений (*мультипликации*), поскольку групповой метод производства предполагает изготовление множества однотипных структур (диодов или транзисторов) на одной пластине. Мультипликация осуществляется путем шагового перемещения фотопластинки и последовательного экспонирования на нее промежуточного изображения с одновременным уменьшением его размеров до окончательного масштаба. Полученный фотошаблон хранится в качестве *эталона*; с него переснимаются рабочие копии ¹.

Одиночные фотошаблоны являются редким исключением. Обычно для изготовления любого диода или транзистора требуется комплект согласованных шаблонов (из 4—5 шт. и более), каждый из которых используется на соответствующем этапе технологического процесса для создания той или иной конфигурации «окон» в оксидной маске.

Глава пятая

РАЗНОВИДНОСТИ ТРАНЗИСТОРОВ

5-1. ТОЧЕЧНЫЙ ТРАНЗИСТОР

Точечные транзисторы были первыми полупроводниковыми трехполосниками, в которых обнаружился эффект усиления.

Открытие усилительного эффекта в полупроводниковом приборе было предопределено интенсивными исследованиями германиевых и кремниевых точечных диодов в 1943—1948 гг., поскольку эти диоды играли важную роль в новой для того времени области науки и техники — радиолокации.

В июне 1948 г. Д. Бардин и В. Браттейн, исследуя поверхностные явления в точечном германиевом диоде, расположили вблизи основного контакта диода еще одну иглу — зонд (рис. 5-1). Оказалось, что если зонд находится под достаточно большим обратным напряжением, то ток в его цепи повторяет изменения тока через основной контакт, а мощность в цепи зонда может превышать мощность в цепи контакта [84]. Это явление после соответствующих технологических и конструктивных разработок было реализовано в виде точечных транзисторов. Теория последних достаточно сложная; поэтому ограничимся описанием ее основных положений.

При формовке контактов точечного транзистора ² под эмиттерной иглой образуется слой *p*-типа, а под коллекторной иглой — *d в о й н о й p-n* слой (рис. 5-2). Такая «асимметрия» контактов обу-

¹ В процессе фотолитографии фотошаблоны изнашиваются, и их время от времени приходится заменять.

² Далее везде имеются в виду германиевые транзисторы, так как в период развития точечных транзисторов методы очистки кремния и получения в нем *p-n* переходов еще не были освоены.

словлена как различием в материале игл, так и различием в режиме формовки. В результате точечный транзистор оказывается четырехслойным трехпереходным прибором, и это наряду с полусферической структурой контактов объясняет существенные отличия его от плоскостного транзистора p - n - p . Главным из этих отличий является то, что у точечных транзисторов имеет место усиление тока в схеме с общей базой, т. е. $\alpha > 1$. Как правило, $\alpha = 2 \div 5$, но это значение может доходить до 10—20 и больше. Явления, лежащие в основе этой особенности, рассмотрены в первых трех изданиях данной книги.

Поскольку тепловой ток не влияет на значение α , положим $I_{k0} = 0$ и будем считать, что эмиттерному току $I_э$ соответствует в цепи коллектора ток $I_к$ (рис. 5-3). Электронная составляющая

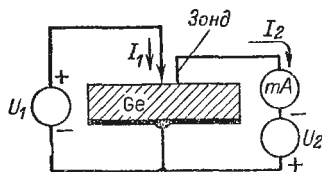


Рис. 5-1. Схема опыта, в котором были обнаружены усиленные свойства полупроводникового прибора.

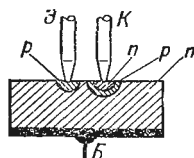


Рис. 5-2. Структура точечного транзистора.

тока $I_к$ в сечении Π_3 равна $\gamma_3 I_к$, где γ_3 — коэффициент инжекции для перехода Π_3 (если рассматривать слой n_2 как эмиттер электронов). Электронный ток через переход Π_2 будет равен $\kappa_{3-2} (\gamma_3 I_к)$, где κ_{3-2} — коэффициент переноса электронов от перехода Π_3 к переходу Π_2 . Дырочный ток через переход Π_2 равен $\alpha_{1-2} I_э$, где α_{1-2} — интегральный коэффициент передачи тока от перехода Π_1 к переходу Π_2 . Очевидно, что сумма электронного и дырочного токов в сечении Π_2 должна быть равна току коллектора, т. е.

$$\alpha_{1-2} I_э + \kappa_{3-2} \gamma_3 I_к = I_к.$$

Отсюда для транзистора в целом получаем интегральный коэффициент передачи тока в следующем виде:

$$\alpha = \frac{I_к}{I_э} = \frac{\alpha_{1-2}}{1 - \alpha_{3-2}}.$$

Здесь $\alpha_{3-2} = \kappa_{3-2} \gamma_3$ — интегральный коэффициент передачи тока от перехода Π_3 к переходу Π_2 .

Из полученного выражения видно, что если κ_{3-2} и γ_3 близки к единице, то коэффициент α может быть весьма большим, несмотря на то что $\alpha_{1-2} < 1$. Условие $\alpha_{3-2} \approx 1$ применительно к транзистору n_2 - p_2 - n_1 требует, чтобы его база p_2 была достаточно тонкой и имела большее удельное сопротивление по сравнению с «эмиттером» n_2 . Оба эти условия соблюдаются в точечном транзисторе.

На рис. 5-4 показаны коллекторные характеристики точечного транзистора в схеме ОБ. По форме они напоминают характеристики плоскостного транзистора в схеме ОЭ (см. рис. 4-21, а). Это вполне естественно, поскольку участок $n_1-p_2-n_2$ действительно включен по схеме ОЭ с оборванной базой.

Зависимость коэффициента α от напряжения U_K и тока I_3 объясняется теми же факторами, что и в плоскостном транзисторе (см. § 4-6). Обычно $\alpha > 1$ при $U_K = 5 \div 10$ В. При меньших напряжениях, когда $\alpha < 1$, точечный транзистор в принципе подобен плоскостному, хотя и имеет значительные количественные отличия: допустимая мощность не превышает 100 мВт, тепловой ток I_{K0} составляет

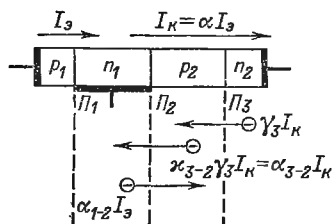


Рис. 5-3. Структура тока через коллекторный переход точечного транзистора.

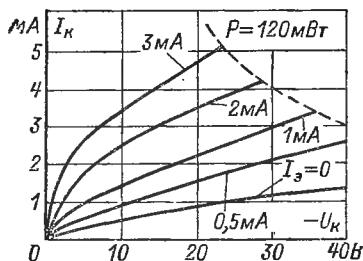


Рис. 5-4. Семейство коллекторных характеристик точечного транзистора.

1—2 мА, коллекторное сопротивление r_K лежит в пределах 7—15 кОм, сопротивления r_3 и r_6 доходят до 500 Ом и больше. Таким образом, параметры точечных транзисторов в общем значительно хуже, чем у плоскостных (не говоря о большем разбросе параметров, высоком уровне шумов, низкой стабильности и других недостатках).

Что касается главной качественной особенности точечных транзисторов — большого значения α , то она привлекательна не возможностью усиливать ток (плоскостной транзистор в схеме ОЭ дает большее усиление), а главным образом возможностью получать характеристики с участком отрицательного сопротивления [85]. Разработка плоскостных приборов, обладающих аналогичными характеристиками (см. ниже), привела к окончательному «отмиранию» точечных транзисторов.

5-2. ЛАВИННЫЙ ТРАНЗИСТОР

Лавинные транзисторы по сравнению с обычными транзисторами не имеют существенных структурных или конструктивных особенностей. Название «лавинный транзистор» относится в основном к их специфическому режиму работы, когда используется

ударная ионизация в коллекторном переходе и в результате получаются S-образные (неоднозначные по напряжению) выходные характеристики при включении ОЭ (рис. 5-5).

В § 4-7 семейство $I_K(U_K)$ для схемы ОЭ (см. рис. 4-21) было получено в условиях заданного тока базы, когда эмиттерный переход транзистора был все время открыт. При этом семейство располагалось в интервале коллекторных напряжений от 0 до U_B , в пределах которого $\alpha < 1$. Две кривые такого семейства (при $I_B > 0$ и $I_B = 0$) воспроизведены на рис. 5-5 пунктиром. S-образные кривые получают лишь тогда, когда эмиттерный переход закрыт при напряжениях $U_K < U_B$ и открывается при $U_K > U_B$. Поясним происхождение участка с отрицательным сопротивлением с помощью схемы на рис. 5-5, а.

Сначала положим, что $R_6 = 0$ и, следовательно, запирающая э. д. с. E_6 приложена непосредственно к эмиттерному переходу. Будем считать, что при этом эмиттерный ток отсутствует¹. Тогда ток базы отрицателен (втекает в базу) и равен коллекторному току. Если коллекторное напряжение U_K положительно, но его значение меньше E_6 , то разность потенциалов U_{K6} будет отрицательной и в цепи база — коллектор будет протекать ток I_{K0} . Этот ток сохранится при $U_K = 0$ и при небольших отрицательных значениях U_K (рис. 5-5, б). Однако при больших отрицательных значениях U_K , когда разность потенциалов U_{K6} делается сравнимой со значением U_M , заметную роль начинает играть ударная ионизация в коллекторном переходе и соответственно токи базы и коллектора возрастают в M раз [см. (2-55)]. В результате характеристика $I_K(U_K)$, показанная на рис. 5-5, б пунктиром, совпадает с кривой $I(U)$ на рис. 2-28.

Пусть теперь последовательно с E_6 включен резистор R_6 (рис. 5-5, а). Тогда возрастание отрицательного базового тока в процессе ударной ионизации будет сопровождаться уменьшением запирающего напряжения на эмиттерном переходе. При некотором достаточно большом токе это напряжение делается равным нулю и открывается эмиттерный переход (на рис. 5-5, б — в точке а). В дальнейшем напряжение U_{63} меняется слабо, поскольку эмиттерный переход уже открыт и имеет малое сопротивление. Соответственно базовый ток остается почти постоянным и равным E_6/R_6 . Постоянство базового тока означает, что приращения ΔI_K и ΔI_3 одинаковы, а следовательно, дифференциальный коэффициент передачи эмиттерного тока равен единице. С учетом ударной ионизации это условие можно

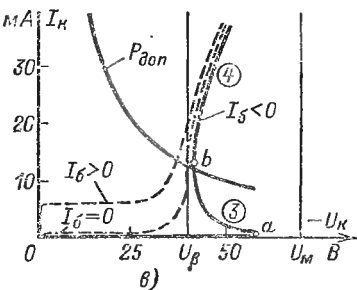
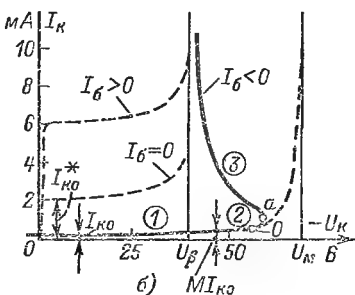
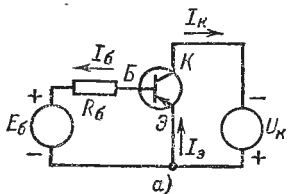


Рис. 5-5. Лавинный транзистор, включенный по схеме ОЭ (а), и его выходные характеристики в области малых (б) и больших (в) токов.

¹ На самом деле в режиме отсечки имеется остаточный эмиттерный ток, но он значительно меньше тока коллектора, см. § 15-2.

$$\alpha_M = M\alpha = 1. \quad (5-1)$$

Главное обстоятельство, на котором основаны последующие выводы, состоит в том, что α является функцией эмиттерного тока (см. § 4-6). Эта функция показана на рис. 5-6.

После отпирания эмиттерного перехода α возрастает, а α_M остается равным единице. Это означает, что коэффициент ударной ионизации M должен уменьшаться. Соответственно должно уменьшаться коллекторное напряжение. Иначе говоря, при токах, превышающих значения I_a , увеличение тока сопровождается уменьшением напряжения, т. е. дифференциальное сопротивление оказывается отрицательным.

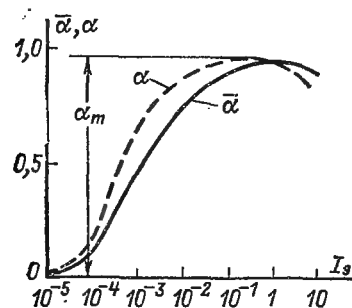


Рис. 5-6. Зависимость коэффициентов передачи $\bar{\alpha}$ и α от эмиттерного тока (единицы тока — условные).

Из всего сказанного следует, что наличие «отрицательного» участка характеристики обусловлено сопротивлением R_6 (в отсутствие внешнего резистора его роль в принципе может играть собственное сопротивление базы). Естественно, что «отрицательный» участок будет выражен тем ярче, чем больше разность $U_a - U_b$; напряжение U_a в свою очередь тем больше, чем больше ток I_a , т. е. чем больше E_6 и чем меньше R_6 . Что касается величины отрицательного сопротивления, то она растет (а значит, S-образность характеристики делается более выраженной) при резкой зависимости $\alpha(I_a)$ в области малых эмиттерных токов¹.

Теперь проведем элементарный анализ лавинного транзистора. В основу анализа положим выражение (4-8), умножив его правую часть на коэффициент ударной ионизации²:

$$I_K = M(\alpha I_a + I_{K0}). \quad (5-2)$$

Подставляя $I_a = I_K + I_6$, легко получить другую запись:

$$I_K = \frac{(M\alpha) I_6 + M I_{K0}}{1 - (M\alpha)}. \quad (5-3)$$

Коэффициент M является функцией напряжения, а коэффициент α — функцией тока³. Поэтому в целом выражение (5-2) пред-

¹ Кроме того, наклон характеристики зависит от функции $M(U_K)$, т. е. согласно (2-55) от показателя n .

² В (5-2) и последующих формулах под α подразумевается интегральный коэффициент $\bar{\alpha}$ (см. с. 188). Однако для простоты черточка над символом не ставится.

³ Строго говоря, α является также и функцией напряжения (см. § 4-6), но в данном случае зависимость $\alpha(U_K)$ менее существенна, чем $M(U_K)$, и учет ее только усложняет анализ, не внося принципиальных уточнений.

ставляет собой вольт-амперную характеристику лавинного транзистора в неявном виде. Поскольку кривая неоднозначна по напряжению, следует считать заданной величиной ток. Соответственно формулу (5-2) целесообразно преобразовать к виду $U_k(I_k)$, используя для коэффициента M выражение (2-55). Тогда после преобразования получим:

$$U_k = U_M \sqrt[n]{1 - \frac{\alpha I_s + I_{k0}}{I_k}}, \quad (5-4)$$

где $I_s = I_k + I_0$ (в нашем случае $I_0 < 0$). Функцию (5-4) можно строить по точкам, задавая ток I_k и определяя соответствующие ток I_s и коэффициент $\alpha(I_s)$.

На начальном участке 1 (см. рис. 5-5, б) эмиттерный переход заперт и, следовательно, $\alpha = 0$. Поэтому согласно (5-4) напряжение U_k растет с увеличением тока I_k . В точке 0, в которой отпирается эмиттерный переход, коллекторный ток равен (по модулю) току базы:

$$I_0 = \frac{E_0}{R_0}. \quad (5-5)$$

При этом из (5-4) получаем напряжение

$$U_0 = U_M \sqrt[n]{1 - \frac{I_{k0} R_0}{E_0}}. \quad (5-6)$$

Сравнивая (5-6) и (4-77), приходим к выводу, что неравенство $U_0 > U_\beta$ легко выполняется при условии

$$\frac{E_0}{R_0} > \frac{I_{k0}}{\alpha_0}. \quad (5-7)$$

Здесь α_0 — номинальный коэффициент передачи, который подразумевался в формуле (4-77). Условие (5-7) должно выполняться при максимальном тепловом токе.

На переходном участке 2 рост напряжения замедляется и дифференциальное сопротивление делается все меньше. Это следует из выражения (5-4), в котором увеличение тока I_k сопровождается увеличением слагаемого αI_s в числителе дроби. В критической точке a дифференциальное сопротивление обращается в нуль, а затем делается отрицательным.

Анализ показывает, что координаты точки a близки к координатам точки 0 и их можно определять по формулам (5-5) и (5-6).

Отрицательное сопротивление на участке 3 легко получить, продифференцировав (5-4) по току. Однако в общем случае получается громоздкое выражение. Рассмотрим частный случай, когда ток I_k много больше тока I_s и тем более тока $I_0 \approx I_0$. Тогда можно пренебречь в формуле (5-4) тепловым током I_{k0} и положить $I_k \approx I_s$. При этом вольт-амперная характеристика упрощается и принимает вид:

$$U_k \approx U_M \sqrt[n]{1 - \alpha}. \quad (5-8)$$

Отсюда, продифференцировав U_k по току, получим упрощенную формулу для сопротивления:

$$r_k \approx \frac{U_M \left(-\frac{d\alpha}{dI_g} \right)}{n(1-\alpha)^{(n-1)/n}}. \quad (5-9)$$

С ростом тока производная $d\alpha/dI_g$, а вместе с ней и сопротивление r_k уменьшаются, стремясь к нулю. Если бы функция $\alpha(I_g)$ была монотонной, то кривая $U_k(I_k)$ асимптотически приближалась бы к вертикали U_β , как показано на рис. 5-5, б. В действительности асимптотический процесс не имеет места (рис. 5-5, в), так как при больших токах коэффициент α уменьшается и производная $d\alpha/dI_g$ делается отрицательной. Соответственно сопротивление r_k должно изменить знак в некоторой точке в, вследствие чего на характеристике получится второй положительный участок (участок 4) *.

Мы исследовали одну кривую с параметром E_6 . Изменяя параметр, получим семейство кривых, подобных той, которая изображена на рис. 5-5, в. Такое семейство характерно тем, что точка в смещается сравнительно мало, тогда как точка а смещается в пределах от U_β до U_M .

Хотя лавинные транзисторы впервые описаны еще в 1956 г. [86], их практическое применение за истекшие двадцать лет носит эпизодический характер, главным образом в области формирования наносекундных импульсов [87] **. Основным препятствием для более широкого применения (особенно в интегральных схемах), по-видимому, является то, что отрицательный участок вольт-амперной характеристики получается только при сравнительно высоком коллекторном напряжении ($U_k > U_\beta$), равном обычно 10—15 В и более. В этом отношении они уступают тиристорам, описываемым в следующем разделе, хотя быстродействие тиристоров существенно ниже.

5-3. ТИРИСТОРЫ

Полупроводниковым приборам с тремя p - n переходами, имеющим участок с отрицательным сопротивлением на вольт-амперной характеристике, присвоено общее название *тиристоры*.

* По той же причине вертикаль U_β не является асимптотой и для других кривых, показанных на рис. 5-5, б. Все характеристики, строго говоря, должны стремиться к вертикали U_M . Однако практически это можно заметить только в импульсном режиме, так как в противном случае из-за разогрева транзистора наступает тепловой пробой, маскирующий описанные явления (см. с. 134).

** Пробой перехода не сопровождается инжекцией неосновных носителей, а значит, их накоплением и рассасыванием. Поэтому для лавинных транзисторов характерно весьма высокое быстродействие. Время переключения из режима малых токов в режим больших токов и обратно ограничивается в основном перезарядом барьерной емкости перехода.

Двухэлектродные тиристоры назовем *динисторами*, а трехэлектродные — *тринисторами*¹.

Динистор. Рассмотрим работу диода, состоящего из четырех чередующихся слоев $p_1-n_1-p_2-n_2$ (рис. 5-7, а). Если подать на него не очень большое напряжение U плюсом на слой p_1 и минусом на слой n_2 , то потечет ток, как показано стрелкой. В результате переходы Π_1 и Π_3 будут работать в прямом направлении, а переход Π_2 — в обратном. Таким образом, получится как бы сочетание двух транзисторов в одном приборе (рис. 5-7, б)²: одним транзистором является комбинация слоев $p_1-n_1-p_2$, другим — комбинация слоев $n_1-p_2-n_2$. Слои p_1 и n_2 являются эмиттерами, n_1 и p_2 — базами для одного транзистора и коллекторами для второго. Во избежание путаницы их называют базами. Переход Π_2 называют коллекторным.

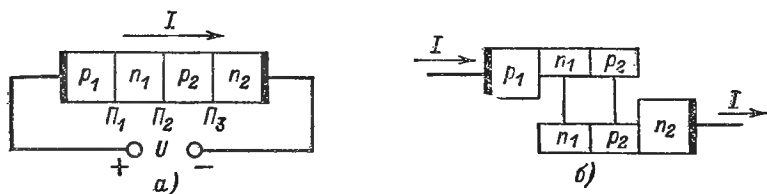


Рис. 5-7. Структура динистора (а) и его двухтранзисторный эквивалент (б).

Пока коллекторный переход работает в обратном направлении, практически все приложенное напряжение U падает на нем. Поэтому при больших значениях U следует учитывать ударную ионизацию в этом переходе. Примем для дырок и электронов один и тот же коэффициент умножения M (чтобы не усложнять выкладки) и обозначим через α_1 и α_3 интегральные коэффициенты передачи тока от переходов Π_1 и Π_3 к переходу Π_2 . Тогда ток последнего можно записать в следующем виде:

$$I_{\Pi 2} = M (\alpha_1 I_{\Pi 1} + \alpha_3 I_{\Pi 3} + I_{k0}), \quad (5-10)$$

где I_{k0} — сумма теплового тока, тока термогенерации и тока утечки в переходе Π_2 .

Поскольку токи через все три перехода одинаковы и равны внешнему току I , легко находим:

$$I = \frac{MI_{k0}}{1 - M\alpha}. \quad (5-11)$$

¹ Вместо этих названий часто используют термины *диод-тиристор* и *триод-тиристор*. Более подробные сведения о приборах подобного типа можно найти в [88—91].

² Комбинация транзисторов $p-n-p$ и $n-p-n$, показанная на рис. 5-7, б, действительно обладает свойствами динистора и может быть использована на практике.

Здесь $\alpha = \alpha_1 + \alpha_3$ — суммарный коэффициент передачи тока от обоих эмиттеров к коллекторному переходу.

Следует заметить, что при изготовлении динистора за основу берется пластина типа p . Через одну из ее поверхностей осуществляется двухкратная диффузия (сначала доноров, потом акцепторов), в результате которой образуются слои n_1 и p_1 . Через другую поверхность осуществляется однократная диффузия доноров и образуется слой n_2 . Поскольку глубины диффузионных слоев не превышают 10 мкм, а толщина исходной пластины составляет 250—300 мкм, одна из баз (n_1) оказывается *тонкой*, а другая (p_2) — *толстой*. В тонкой базе, как и у транзисторов, выполняется условие $w < L$ и максимальное значение коэффициента α_1 (см. рис. 5-6) близко к единице. В толстой базе действительно соотношение $w > L$, и соответственно максимальное значение коэффициента α_3 существенно меньше единицы (обычно 0,1—0,2). Тем не менее сумма *м а л ь н ы х* значений α_1 и α_3 превышает единицу, что необходимо для работы динистора и тиристорov вообще. Обычно $\alpha_m \approx 1,05 \div 1,15$. Наличие толстой базы с ее малым коэффициентом переноса играет в целом положительную роль, так как в области малых токов суммарный коэффициент α нарастает медленнее, а это обеспечивает большие напряжения переключения (см. петит в конце данного раздела).

Выражение (5-11) в неявном виде является вольт-амперной характеристикой динистора, так как параметр M в правой части зависит от напряжения¹. Структура выражения (5-11) такая же, как в случае лавинного транзистора при $I_6 = 0$ [см. (5-3)]. Такое сходство вполне естественно, поскольку оба «составляющих транзистора» в динисторе (рис. 5-7, б) включены по схеме ОЭ с оборванной базой.

Вольт-амперная кривая динистора вместе с его условным обозначением показана на рис. 5-8. Как видим, она подобна характеристике лавинного транзистора в схеме ОЭ (см. рис. 5-5). Однако существенным преимуществом динисторов является то, что рабочее напряжение в области *б о л ь ш и х* токов у них значительно меньше и почти не зависит от тока. Кроме того, динисторы работают без всякого предварительного смещения в цепи базы в отличие от лавинных транзисторов, у которых такое смещение необходимо (рис. 5-5, а). Критические точки характеристики на рис. 5-8, в которых $r = dU/dI = 0$, называют соответственно точкой *прямого переключения* (ПП) и точкой *обратного переключения* (ОП).

Происхождение отрицательного участка на характеристике динистора обусловлено той же причиной, что и в лавинном транзисторе, т. е. тем, что у обоих приборов на этом участке задан постоянный ток базы (у динистора он равен нулю). Поэтому должно выполняться соотношение $dI_K = dI_B$, т. е. *д и ф ф е р е н ц и а л ь н ы й* коэффициент α должен быть все время равен единице. С ростом тока величина α стремится возрасти, но это возрастание предотвращается уменьшением напряжения на коллекторном переходе, т. е. ослаблением ударной ионизации. Такой же вывод следует из формул (5-3) и (5-11), в которых *з п а м е н а т е л ь н ы й* не может

¹ Ток I_{K0} при том его определении, которое дано в формуле (5-10), тоже зависит от напряжения. Однако учет этой зависимости наряду с зависимостью $M(U)$ сильно усложняет задачу. В некоторых случаях (например, если переход P_2 зашунтирован небольшим, заранее известным сопротивлением), можно пренебречь функцией $M(U)$ и считать зависимость от напряжения сосредоточенной в функции $I_{K0}(U)$. Общая методика анализа при этом не меняется.

быть отрицательным, и, следовательно, начиная с некоторой рабочей точки, увеличение интегрального коэффициента α должно сопровождаться уменьшением коэффициента M , т. е. уменьшением коллекторного напряжения.

Несмотря на определенное сходство с лавинным транзистором, динистор имеет принципиальную особенность. Эту особенность легко показать, если представить вольт-амперную характеристику в форме $U(I)$, аналогичной форме (5-4) для лавинного транзистора. Подставив выражение (2-55) в (5-11) и решив последнее относительно напряжения, получим:

$$U = U_M \sqrt[n]{1 - \frac{\alpha I + I_{к0}}{I}}. \quad (5-12)$$

Сравнивая (5-4) и (5-12), приходим к следующему важному выводу. У лавинного транзистора, у которого $\alpha < 1$ при любом токе, напряжение U_k всегда имеет конечную величину. У динистора,

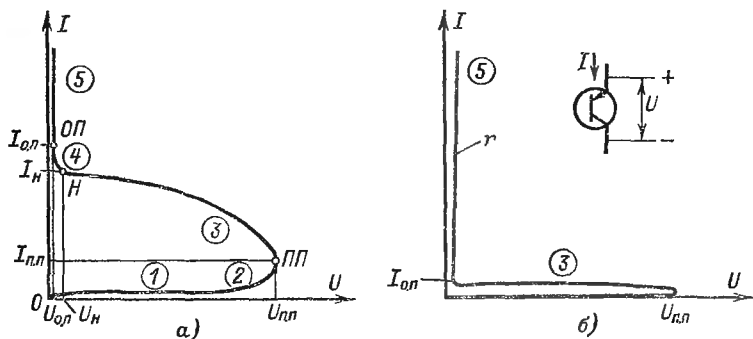


Рис. 5-8. Вольт-амперная характеристика динистора.

а — начальный участок; б — полная кривая.

у которого суммарный коэффициент $\alpha = \alpha_1 + \alpha_3$ может превышать единицу, напряжение U (точнее напряжение на коллекторном переходе) делается равным нулю при некотором конечном токе I . При еще большем токе формулы (5-11) и (5-12) становятся недействительными, так как коллекторный переход оказывается смещенным в прямом направлении и механизм работы динистора качественно изменяется. Рассмотрим отдельные участки характеристики, показанной на рис. 5-8.

Начальный участок 1 характерен очень малыми токами, при которых можно считать $\alpha \approx 0$. Сопротивление на этом участке весьма велико, поэтому заданной величиной всегда бывает напряжение, а ток можно найти по формуле (5-11).

На переходном участке 2 рост напряжения замедляется, а сопротивление резко падает. Эти изменения являются следствием увеличения коэффициента α и могут быть легко оценены с помощью выражения (5-12).

В конце второго участка, в точке $ПП$, сопротивление обращается в нуль, а затем (при заданном токе) становится отрицательным. Координаты точки прямого переключения определяются условием $dU/dI = 0$.

Напряжение $U_{п.п}$ обычно близко к значению U_M и может лежать в широких пределах — от 25—50 до 1000—2000 В *. Ток $I_{п.п}$ лежит в пределах от долей микроампера до нескольких миллиампер в зависимости от площади переходов.

На отрицательном участке 3 характеристика по-прежнему описывается формулой (5-12), которую, однако, можно упростить, полагая $\alpha I \gg I_{к0}$. Тогда

$$U \approx U_M \sqrt[n]{1 - \alpha}, \quad (5-13)$$

где α увеличивается с ростом тока. Дифференцируя (5-13) по току, получаем сопротивление на этом участке:

$$r = - \frac{U_M \left(\frac{d\alpha}{dI} \right)}{n(1 - \alpha)^{(n-1)/n}}. \quad (5-14)$$

Отсюда видно, что величина сопротивления должна существенно меняться с изменением тока. Характер этого изменения определяется функцией $\alpha(I)$, чаще всего сопротивление r возрастает (по модулю) с ростом тока. Среднее значение $|r|$ между точками $ПП$ и $ОП$ лежит обычно в пределах от 5—10 до 50—100 кОм.

Коллекторное напряжение, уменьшаясь на участке 3, делается равным нулю в точке H **. Из формулы (5-12) при $U = 0$ получаем соотношение

$$I = \frac{I_{к0}}{1 - \alpha}, \quad (5-15)$$

из которого определяется ток I_H . Поскольку этот ток несравненно больше, чем $I_{к0}$, его можно определять из условия

$$\alpha = \alpha_1 + \alpha_3 \approx 1, \quad (5-16)$$

пользуясь графиками $\alpha(I)$.

Напряжение U_H является суммой напряжений на эмиттерных переходах, так как $U_{п2} = 0$. Используя формулу (4-7) при $U_k = 0$, $I_s = I_H$ и считая оба эмиттерных перехода одинаковыми, получаем:

$$U_H \approx 2\varphi_T \ln \frac{I_H}{I'_{s0}}. \quad (5-17)$$

Это напряжение составляет несколько десятых долей вольта у германиевых диодисторов и 0,5—1 В у кремниевых.

* Можно изготовить аналогичные приборы с рабочими напряжениями всего в несколько вольт [92].

** Точка H обозначает границу режима насыщения — режима, в котором и эмиттерные, и коллекторный переходы работают в прямом направлении.

При токе $I > I_n$ переход P_2 , будучи смещен в п р я м о м направлении, инжектирует носители навстречу тем потокам, которые поступают от эмиттеров. Инжектируемый компонент тока I_{n2} равен разности между собираемым компонентом $(\alpha_1 I_{n1} + \alpha_3 I_{n3})$ и полным током I_{n2} . Поэтому, если для простоты положить $\alpha_I = 0$ (т. е. считать, что носители, инжектируемые переходом P_2 , не доходят до эмиттеров) и принять условие $U \gg \varphi_T$ для всех трех переходов, то напряжение на открытом диносторе можно выразить с помощью формулы (4-7) в виде суммы напряжений на переходах:

$$U = \varphi_T \left[\ln \frac{I_{n1}}{I_{\infty 01}} - \ln \frac{(\alpha_1 I_{n1} + \alpha_3 I_{n3}) - I_{n2}}{I_{\infty 02}} + \ln \frac{I_{n3}}{I_{\infty 03}} \right] \quad (5-18a)$$

[токи $I'_{\infty 0}$ заменены на $I_{\infty 0}$, так как принято, что $\alpha_I = 0$, см. (4-36)].

Учитывая, что $I_{n1} = I_{n2} = I_{n3} = I$ и полагая токи $I_{\infty 0}$ одинаковыми у всех переходов, получаем простое приближенное выражение:

$$U \approx \varphi_T \ln \frac{I/I_{\infty 0}}{\alpha - 1}. \quad (5-18b)$$

Вблизи точки H , где $\alpha \approx 1$, увеличение тока, а вместе с ним коэффициента α приводит к сильному увеличению разности $\alpha - 1$ и напряжение несколько уменьшается (участок 4). В точке ОП напряжение достигает минимума и в дальнейшем растет с ростом тока (участок 5) из-за падения напряжения в толстой базе p_2 .

Обычно параметры точек H и ОП близки друг к другу, поэтому можно вычислять координаты точки ОП по формулам (5-16) и (5-17).

При отрицательном напряжении U переход P_2 смещен в прямом направлении; дырки инжектируются в слой n_1 , а электроны — в слой p_2 . Переходы P_1 и P_3 смещены в о б р а т н о м направлении и являются в данном случае коллекторными. Динистор в этом режиме эквивалентен двум последовательно включенным транзисторам (p - n - p и n - p - n) с оборванными базами.

Важной проблемой при разработке динисторов и других аналогичных приборов является обеспечение п л а в н о г о изменения коэффициента α в области малых токов [93]. Действительно, как уже отмечалось, 2-й (переходный) участок вольт-амперной кривой (рис. 5-8, а) характерен заметной и растущей ролью слагаемого αI по сравнению с током I_{k0} в формуле (5-12). Значит, чем медленнее увеличивается α с ростом тока, тем позднее (при больших токах) начнется 2-й участок и тем больше будет напряжение переключения, что обычно желательно в таких приборах. С этой точки зрения предпочтительным материалом для динисторов является кремний, так как у кремниевых переходов благодаря большей роли процессов генерации — рекомбинации коэффициент инжекции при малых токах близок к нулю и с ростом тока увеличивается весьма медленно (см. с. 208). Еще одним преимуществом кремния является малое значение тока в запертом состоянии прибора. Однако, с другой стороны, кремниевые переходы характерны большим значением прямого напряжения, что несколько ухудшает параметры динистора в открытом состоянии.

Чтобы ослабить зависимость α (I) при малых токах (особенно у германиевых структур), часто шунтируют эмиттерный переход небольшим сопротивлением R . Тогда часть общего тока ответвляется в это сопротивление, минуя эмиттер, в результате чего эмиттерный ток, а вместе с ним и коэффициент α при прочих равных условиях уменьшаются.

Тринистор. Снабдим базу диностора n_1 внешним выводом и используем этот третий электрод для задания дополнительного тока через переход p_1-n_1 (рис. 5-9) *. Для такого прибора (тринистора) принята та же терминология, что и для обычного транзистора: выходной ток называется коллекторным, а управляющий — базовым. Эмиттером считается слой, примыкающий к базе, хотя с физической точки зрения эмиттером является и второй внешний слой (в нашем случае n_2). Условное обозначение тринистора вместе с семейством характеристик показано на рис. 5-10. Как видим, увеличение управляющего тока I_6 приводит прежде всего к умень-

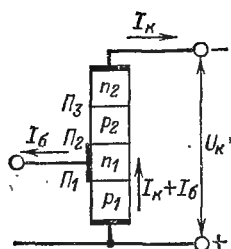


Рис. 5-9. Структура тринистора.

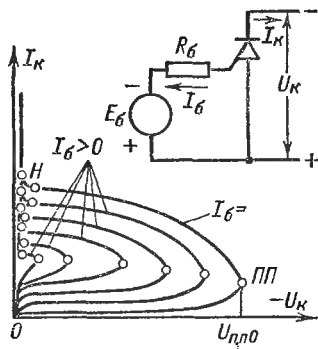


Рис. 5-10. Вольт-амперные характеристики тринистора при положительном токе базы.

шению напряжения прямого переключения. Кроме того, несколько возрастает ток прямого переключения, а ток обратного переключения уменьшается. В результате отдельные кривые с ростом тока I_6 как бы «вписываются» друг в друга вплоть до полного исчезновения отрицательного участка (такую кривую называют *спрямленной характеристикой*).

Элементарный анализ тринистора можно провести, исходя из формулы (5-10), в которой нужно положить $I_{п3} = I_{п2} = I_K$ и $I_{п1} = I_K + I_6$. Тогда вместо формулы (5-11) получим для тока I_K более общее выражение:

$$I_K = \frac{MI_{K0} + (M\alpha_1) I_6}{1 - M\alpha} \quad (5-19)$$

Здесь $\alpha = \alpha_1 + \alpha_3$ — суммарный коэффициент передачи, в котором составляющая α_3 является функцией тока I_K , а составляющая α_1 — функцией суммы токов $I_K + I_6$. Задавая положительный ток I_6 при $I_K = 0$, мы задаем н а ч а л ь н о е значение коэффициен-

* В качестве управляющей используется тонкая база, у которой коэффициент передачи α_1 близок к единице.

та α_1 . Поэтому любому току I_k будет соответствовать большее значение α_1 , а значит, и большее значение α , чем при $I_6 = 0$.

Решая (5-19) относительно M и используя выражение (2-55), нетрудно представить вольт-амперные характеристики тринистора в форме $U_k(I_k)$:

$$U_k = U_M \sqrt[n]{1 - \frac{\alpha I_k + I_{k0} + \alpha_1 I_6}{I_k}}. \quad (5-20)$$

В частном случае, при $I_6 = 0$, получается характеристика динистора (5-12). Выражение (5-20) ясно показывает, что данному току I_k соответствует тем меньшее напряжение U_k , чем больше ток I_6 (рис. 5-10). Рассмотрим отдельные участки этого семейства.

На начальном участке (до точек ПП) мы имеем по существу семейство характеристик обычного транзистора в схеме ОЭ. Это следует из структурного сходства выражений (5-19) и (5-3), если в последнем считать $I_6 > 0$.

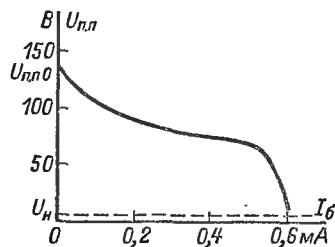


Рис. 5-11. Пусковая характеристика тринистора.

Координаты точек прямого переключения определяются, как и в динисторе, условием $dU_k/dI_k = 0$. Анализ показывает, что ток $I_{н.п}$ возрастает с увеличением тока базы. На рис. 5-11 показана пусковая характеристика тринистора, т. е. зависимость $U_{н.п}(I_6)$.

Координаты точки H , в которой напряжение на коллекторном переходе P_2 падает до нуля, определяются условием $U_k = 0$ в формуле (5-20). Так же как в динисторе, можно в этой точке считать $\alpha \approx 1$ и определять ток I_n из условия

$$\alpha = \alpha_1(I_n + I_6) + \alpha_3(I_n) = 1. \quad (5-21)$$

Отсюда видно, что увеличение тока I_6 , а следовательно и коэффициента α_1 сопровождается уменьшением коэффициента α_3 , а значит, и тока I_n . Соответственно несколько меньше будет и ток $I_{о.п}$ в точке обратного переключения.

Параметры тринистора в открытом состоянии практически не отличаются от параметров динистора, поскольку ток I_k в этой области значительно больше тока I_6 , т. е. токи обоих крайних переходов почти одинаковы.

До сих пор рассматривались кривые с параметром $I_6 > 0$. При этом подразумевалось, что источник базового тока представляет собой э. д. с. $E_6 < 0$, включенную последовательно с сопротивлением R_6 (см. рис. 5-10). В частном случае, при $I_6 = 0$, можно было считать $E_6 = 0$; $R_6 = \infty$. Теперь рассмотрим работу тринистора в условиях обратного смещения ($E_6 > 0$) (рис. 5-12).

Пусть э. д. с. E_6 достаточно велика и эмиттерный переход заперт. Тогда тринистор превращается в транзистор $n_1-p_2-n_2$ (с оборванной базой p_2),

который включен последовательно с сопротивлением R_6 и питается напряжением $E_6 + U_K$. Коллекторный ток при таком включении будет током транзистора в схеме ОЭ с оборванной базой [ср. с (4-70)]:

$$I_K = \frac{MI_{K0}}{1 - M\alpha_3},$$

где α_3 — коэффициент передачи тока от перехода P_3 к переходу P_2 . Реальное запирающее смещение на эмиттерном переходе будет меньше, чем э. д. с. E_6 , на величину $I_K R_6$. С ростом тока I_K смещение будет уменьшаться и при некотором токе I_0 , когда $E_6 - I_0 R_6 = 0$, эмиттерный переход отперется. После этого базовый ток будет иметь неизменную отрицательную величину:

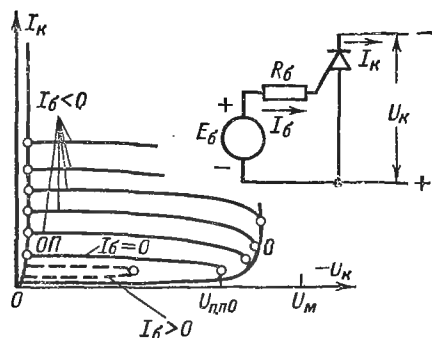
$$I_6 = -I_0 = -\frac{E_6}{R_6}, \quad (5-22)$$

которую можно считать параметром соответствующей характеристики.

Если в формуле (5-20) положить $\alpha_1 = 0$, $\alpha = \alpha_3$ и подставить $I_K = I_0$, можно получить напряжение отпирания эмиттерного перехода:

$$U_0 = U_M \sqrt[n]{1 - \frac{\alpha_3 I_0 + I_{K0}}{I_0}}. \quad (5-23)$$

Рис. 5-12. Вольт-амперные характеристики тринистора при отрицательном токе базы.



Из формулы (5-22) очевидно, что ток отпирания I_0 , будучи равен параметру I_6 , возрастает вместе с модулем параметра. Напряжение отпирания U_0 также несколько увеличивается.

Ток обратного переключения можно найти из уравнения (5-21), считая $I_n \approx I_{0.n}$. В случае малых отрицательных токов базы ток $I_{0.n}$ заметно больше тока I_0 . При больших токах $|I_6|$ эта разница уменьшается. Отношение $I_{0.n}/|I_6|$ можно назвать *коэффициентом усиления при выключении*; согласно [88] он определяется значением $\alpha_1/(\alpha - 1)$ и в обычных тринисторах не превышает 5—10. Очевидно, что с точки зрения управляемости при запираии суммарный коэффициент передачи α не следует делать намного большим единицы¹.

Основная тенденция при разработке современных тиристоров состоит в повышении рабочих токов и напряжений с тем, чтобы полностью заменить соответствующие газоразрядные приборы (газотроны, тиратроны, игнитроны и др.). В настоящее время рабочие токи тринисторов доходят до 1000—5000 А, а рабочие напряжения — до 2—5 кВ. При прочих равных условиях тиристоры значительно превосходят газоразрядные приборы по коэффициенту полезного действия и сроку службы при меньших габаритах и массе.

¹ В этом отношении положительную роль играет наличие в структуре тринистора толстой базы (см. петит на с. 276).

Мощные транзисторы используются в качестве контакторов, коммутаторов тока, а также в преобразователях постоянного напряжения, инверторах и выпрямительных схемах с регулируемым выходным напряжением.

Времена переключения у транзисторов значительно меньше, чем у тиратронов. Даже у мощных приборов с токами в сотни ампер время прямого переключения составляет около 1 мкс, а время обратного переключения не превышает 10—20 мкс. Следует заметить, что наряду с конечной длительностью фронтов напряжения и тока имеют место задержки и фронтов по отношению к моменту подачи управляющего импульса.

Кроме мощных тиристоров разработаны и маломощные высокочастотные варианты. В таких приборах время прямого переключения может составлять десятки, а время обратного переключения — сотни наносекунд. Столь высокое быстродействие обеспечивается малой толщиной слоев и наличием электрического поля в толстой базе. Маломощные быстродействующие тиристоры используются в импульсных схемах, в том числе интегральных.

5-4. УНИПОЛЯРНЫЕ (ПОЛЕВЫЕ) ТРАНЗИСТОРЫ

Униполярными называют такие транзисторы, работа которых основана на использовании основных носителей: только дырок или только электронов. С этой точки зрения обычные транзисторы, рассмотренные ранее, можно назвать *биполярными*, так как в них важную роль играют оба типа носителей: инжекция неосновных носителей одного знака сопровождается компенсацией образующегося заряда основными носителями другого знака.

Второй термин — «полевые транзисторы» характеризует механизм управления током: с помощью электрического поля (а не тока, как в биполярных транзисторах). В этом отношении униполярные транзисторы имеют много общего с электронными лампами. В английской литературе униполярные (полевые) транзисторы носят название FET (Field Effect Transistors).

Униполярные полевые транзисторы имеют две основные разновидности. Одна из них, предложенная в 1952 г. [94], основана на использовании поля в p - n переходе; такие транзисторы мы будем называть *унитронами*¹. Вторая разновидность, предложенная в 1963 г. [95], основана на использовании поля в диэлектрике, расположенном между пластиной полупроводника и металлической пленкой; такие транзисторы со структурой «металл—диэлектрик—полупроводник» называются *МДП транзисторами* (по начальным буквам компонентов структуры).

Унитрон. На рис. 5-13, а показана упрощенная структура унитрона с типичной плоскопараллельной конструкцией. Как

¹ В отечественной литературе унитроны обычно называют «полевыми транзисторами с p - n переходом затвора», что несколько длинно и неудобно.

видим, унитрон представляет собой пластинку полупроводника типа n^1 , у которой на торцах имеются омические контакты, а на одной из граней — слой типа p . Последний образует с пластинкой p - n переход. Электрод, связанный со слоем p , называется *затвором*. Двумя другими электродами являются омические контакты. К ним подключается источник питания. При этом тот контакт, от которого движутся основные носители (в нашем случае — электроны), называется *исток*, а тот, к которому они движутся, — *сток*. Следовательно, в унитроне типа n сток имеет положительную полярность относительно истока; p - n переход работает в обратном направлении, для этого на затвор подается отрицательное смещение относительно истока, как показано на рис. 5-13,б.

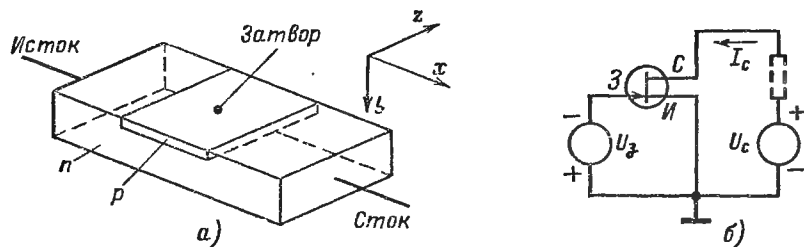


Рис. 5-13. Конструкция унитрона (а) и схема его включения (б).

Принцип действия унитрона очень прост и заключается в том, что при изменении потенциала затвора меняется ширина p - n перехода, а значит, и рабочее сечение пластины. В результате меняются ее сопротивление и соответственно ток в рабочей цепи. Поскольку p - n переход работает в обратном включении, его сопротивление для входного сигнала велико и входная мощность мала. Полезная мощность, определяемая значением питающего напряжения и соотношением сопротивлений пластины и нагрузки, может значительно превышать входную мощность. Таким образом, унитрон является усилительным прибором типа управляемого активного сопротивления, причем он имеет много общего с электронной лампой. Это сходство выражается не только в большом входном сопротивлении, но и в том, что при достаточном отрицательном смещении затвора расширившийся переход может перекрыть все сечение пластинки; это вызовет отсечку тока в рабочей цепи — явление, аналогичное запиранию лампы. Перейдем к более детальному анализу.

Назовем *каналом* рабочую (переменную) часть объема пластины, расположенную под p - n переходом ². Пренебрегая пока «холостыми» участками пластины, прилегающими к истоку и стоку, можно представить структуру унитрона в еще более упрощенном виде (рис. 5-14). Обозначим максимальную толщину канала через a , его ширину через Z и длину через L . Пусть $U_c = 0$, тогда канал будет эквипотенциальным слоем и напряжение на p - n переходе

¹ Вариант типа p , не имея принципиальных отличий, уступает типу n по частотным свойствам, шумам и стабильности [96].

² От термина «канал» происходит второе название унитрона и его вариантов: *канальные транзисторы*.

будет равно U_3 на протяжении всего канала¹. Соответственно в любой точке ширина перехода равна l , а толщина канала

$$w = a - l.$$

Подставляя сюда выражение (2-12) для ширины l и полагая $qN_d = qn = 1/\rho\mu_n$ [см. (1-366)], получаем:

$$w = a - \sqrt{2\varepsilon_0 e \rho \mu_n U_3}.$$

Из условия $w = 0$ легко найти напряжение отсечки:

$$U_{30} = \frac{a^2}{2\varepsilon_0 e \rho \mu_n}. \quad (5-24)$$

Например, при $a \approx 2$ мкм и $\rho = 2$ Ом·см для кремния получается $U_{30} = 6,4$ В. Используя (5-24), запишем толщину канала в более компактной форме:

$$w = a \left(1 - \sqrt{\frac{U_3}{U_{30}}}\right). \quad (5-25)$$

В рабочем режиме, когда $U_c \neq 0$, канал не является эквипотенциальным слоем; в разных точках x по-

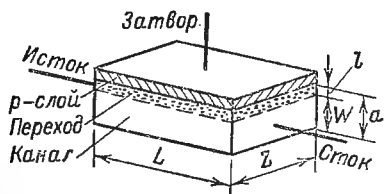


Рис. 5-14. Рабочая часть унитрона.

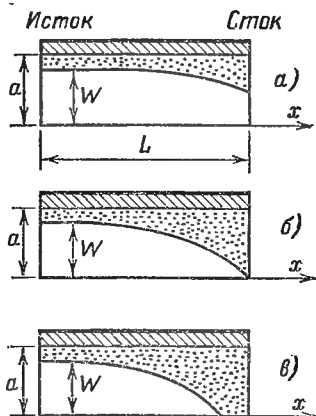


Рис. 5-15. Сечение канала при ненасыщенном режиме унитрона (а), на границе насыщения (б) и при насыщенном режиме (в).

тенциал различен: он меняется от $U_x = 0$ около истока до $U_x = +U_c$ около стока. Поэтому обратное напряжение на $p-n$ переходе, равное $U_3 + U_x$, увеличивается в направлении от истока к стоку. Соответственно ширина перехода в этом направлении растет, а канал сужается (рис. 5-15, а).

В наиболее узком месте (около стока) суммарное напряжение на переходе равно $U_3 + U_c$. С ростом U_c суммарное напряжение в конце концов делается равным U_{30} и переход почти смыкается (рис. 5-15, б). В отличие от случая $U_3 = U_{30}$ это, разумеется, не приводит к отсечке тока, так как само «смыкание» является следствием увеличения тока. Вместо отсечки тока происходит отсечка его приращений, т. е. резкое возрастание диф-

¹ Далее под U_3 будет пониматься модуль напряжения на затворе.

ференциального сопротивления канала. При этом на кривой $I_c(U_c)$, начиная с некоторой точки H , получается почти горизонтальный участок (рис. 5-16, а). Такой режим можно назвать *насыщенным*, а напряжение $U_{c,н}$, при котором он наступает, — *напряжением насыщения*. Эта величина получается из условия $U_3 + U_c = U_{30}$ и равна:

$$U_{c,н} = U_{30} - U_3. \quad (5-26)$$

Напряжение $U_{c,н}$ уменьшается с ростом U_3 .

В режиме насыщения, когда $U_c > U_{c,н}$ потенциал «горловины» канала сохраняет значение $U_{c,н}$ (в противном случае канал

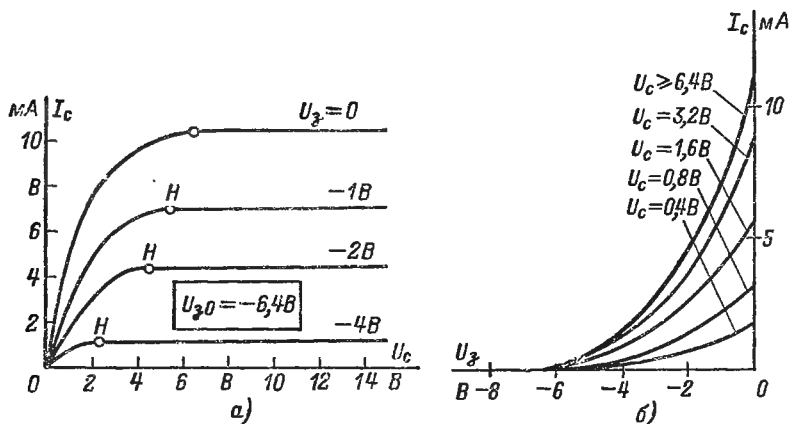


Рис. 5-16. Семейства выходных (а) и передаточных (б) характеристик унитрона.

должен был бы еще больше сужаться, что невозможно), но «горловина» сдвигается относительно стока (рис. 5-15, в). Разность потенциалов $U_c - U_{c,н}$ падает на участке между стоком и «горловиной», а протяженность этого участка определяется формулой (2-12), если положить $U = U_c - U_{c,н}$. Таким образом, в режиме насыщения происходит модуляция длины канала по аналогии с эффектом Эрли в биполярных транзисторах (см. § 4-2).

При расчете характеристик унитрона следует учесть, что сопротивление канала меняется вдоль оси x , поскольку меняется толщина ω . Функцию $\omega(x)$ получаем, заменяя U_3 на $U_3 + U_x$ в формуле (5-25). Тогда падение напряжения на элементарном участке dx составляет:

$$dU_x = I_c dR_x = I_c \frac{\rho dx}{Za} \left(1 - \sqrt{\frac{U_3 + U_x}{U_{30}}} \right)^{-1}.$$

Ток I_c остается неизменным в любом сечении. Поэтому, разделяя переменные и интегрируя обе части соответственно в пределах от 0 до U_c и от 0 до L , можно представить искомую функцию

$I_c(U_c)$ в следующем виде:

$$I_c = \frac{1}{R_{k0}} \left[U_c + \frac{2}{3} \frac{U_3^{3/2} - (U_3 + U_c)^{3/2}}{U_{30}^{1/2}} \right]. \quad (5-27)$$

Здесь через R_{k0} обозначено минимальное дифференциальное сопротивление канала (т. е. $\partial U_c / \partial I_c$ при $U_3 = 0$ и $U_c = 0$):

$$R_{k0} = \frac{\rho L}{aZ}. \quad (5-28)$$

Если $L = 20$ мкм; $a = 2$ мкм; $Z = 1$ мм и $\rho = 2$ Ом·см, то $R_{k0} \approx 200$ Ом.

Семейства характеристик $I_c(U_c)$ и $I_c(U_3)$ показаны на рис. 5-16. Ток насыщения получается подстановкой $U_{c,n}$ из (5-26) в (5-27):

$$I_{c,n} = \frac{1}{R_{k0}} \left[\frac{1}{3} U_{30} - U_3 \left(1 - \frac{2}{3} \sqrt{\frac{U_3}{U_{30}}} \right) \right]. \quad (5-29)$$

Поскольку участок насыщения является основным рабочим участком, определим крутизну S именно в этой области. Дифференцируя (5-29) по U_3 , получаем¹:

$$S = - \frac{\partial I_{c,n}}{\partial U_3} = \frac{1}{R_{k0}} \left(1 - \sqrt{\frac{U_3}{U_{30}}} \right). \quad (5-30)$$

Для $R_{k0} = 200$ Ом и $U_3 = 0,5 U_{30}$ находим: $S \approx 1,5$ мА/В.

Дифференциальное сопротивление r_c в области насыщения согласно (5-29) равно бесконечности. Практически оно может составлять несколько мегаом. Конечная величина r_c обусловлена модуляцией длины канала (см. выше): с ростом U_c длина L уменьшается, уменьшается сопротивлением R_{k0} [см. (5-28)] и ток I_c согласно (5-29) несколько возрастает. В ненасыщенном режиме величины r_c и S легко найти из (5-27).

Переходя к оценке быстродействия унитрона, напомним, что его работа не связана с инжекцией неосновных носителей и распространением их до коллектора. Инерционность унитрона обусловлена зарядом барьерной емкости перехода. Оценим влияние этого фактора.

Емкость перехода заряжается через сопротивление канала. При этом разные участки емкости заряжаются через разные сопротивления в зависимости от расстояния данного участка от истока (рис. 5-17). Чтобы не усложнять анализа, примем, что канал имеет одинаковое сечение Zw на всем протяжении². Тогда емкость p - n пе-

¹ Знак минус в определении крутизны обусловлен тем, что под U_3 понимается абсолютная величина (см. сноску на с. 285). Заметим, что крутизна при нулевом напряжении на затворе равна максимальной проводимости канала: $S_0 = 1/R_{k0}$.

² Такое предположение оправдано тем, что распределение потенциала U_x в канале близко к линейному. Соответственно толщина канала $w(x)$ согласно (5-25) и рис. 5-15 меняется слабо, за исключением области, прилегающей к стоку.

рехода и среднее сопротивление канала будут иметь следующие значения:

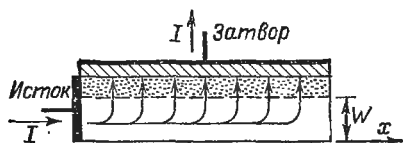
$$C_3 = \frac{\epsilon_0 \epsilon (LZ)}{(a-w)}; \quad (5-31a)$$

$$\bar{R}_k = \frac{0,5\rho L}{Zw}. \quad (5-31b)$$

Постоянная времени затвора будет равна произведению этих параметров:

$$\tau_3 = \bar{R}_k C_3 = \frac{0,5\epsilon_0 \epsilon \rho L^2}{w(a-w)}. \quad (5-32a)$$

Постоянная времени τ_3 зависит от толщины w , которая согласно (5-25) является функцией напряжения U_3 . Минимальное значение τ_3 получается при условии $w = a/2$, т. е. при $U_3 = U_{30}/4$:



$$\tau_{3,\min} = 2\epsilon_0 \epsilon \rho \frac{L^2}{a^2}. \quad (5-32b)$$

Рис. 5-17. Процесс заряда емкости затвора в унитроне.

Для ранее использованных значений ρ , L и a получаем $\tau_{3,\min} \approx 0,5$ нс.

В заключение заметим, что в режиме насыщения вольт-амперная характеристика (5-29) хорошо аппроксимируется более простой квадратичной характеристикой [97]:

$$I_c = \frac{1}{R_{k0}} \frac{(U_{30} - U_3)^2}{3U_{30}}. \quad (5-33a)$$

Дифференцируя (5-33a), получаем выражение для крутизны:

$$S = \frac{2/3}{R_{k0}} \frac{U_{30} - U_3}{U_{30}}. \quad (5-33b)$$

Разновидности унитронов. Плоскостные унитроны сами по себе имеют несколько конструктивно-технологических вариантов (например, слой затвора можно получать путем сплавления, диффузии или эпитаксии). Однако известны и унитроны не плоскостного типа. К ним относятся *текнетрон* и *алкатрон*.

Текнетрон¹ характерен цилиндрической конструкцией (рис. 5-18) с вытекающими отсюда количественными особенностями. Такая конструкция униполярного транзистора была предложена в 1958 г.

Анализ текнетрона выполняется в той же последовательности, что и анализ унитрона. Если обозначить через a максимальный диаметр канала, то в результате анализа получают следующие основные выражения (см. [98], а также 2-е изд. данной книги).

¹ Термин «текнетрон» происходит от первых двух букв фамилии автора (С. Тезнера) и четырех букв (CNET) сокращенного названия фирмы, в которой была проведена разработка.

Вольт-амперная (стоковая) характеристика имеет вид:

$$I_c = \frac{1}{R_{к0}} \frac{U_c \left[(U_{з0} - U_з) - \frac{1}{2} U_c \right]}{U_{з0}}. \quad (5-34)$$

Подставляя сюда напряжение насыщения из (5-26), получаем:

$$I_{с.н} = \frac{1}{R_{к0}} \frac{(U_{з0} - U_з)^2}{2U_{з0}}. \quad (5-35)$$

Крутизна характеристик на рабочем участке

$$S = - \frac{\partial I_{с.н}}{\partial U_з} = \frac{1}{R_{к0}} \left(1 - \frac{U_з}{U_{з0}} \right). \quad (5-36)$$

В приведенных выражениях минимальное сопротивление канала

$$R_{к0} = \frac{\rho L}{\pi a^2 / 4}. \quad (5-37)$$

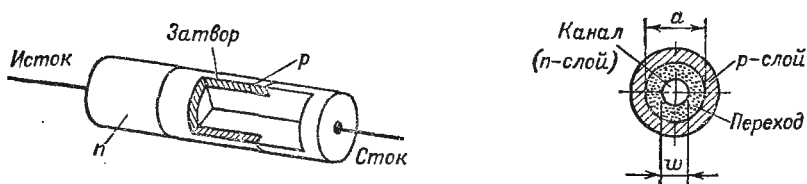


Рис. 5-18. Упрощенная конструкция текнетрона.

Это есть сопротивление цилиндра с диаметром a и высотой L . Напряжение отсечки выражается следующим образом:

$$U_{з0} = \frac{a^2}{16 \epsilon_0 \epsilon_r \mu_n}. \quad (5-38)$$

В текнетроне в отличие от унитрона напряжение затвора $U_з$ меняет сечение канала «со всех сторон». Отсюда следует, что сечение канала при насыщении стягивается не в линию, как у унитрона, а в точку, и следовательно, рассеяние выделяемой мощности более ограниченное.

Значения емкости затвора и среднего сопротивления канала в оптимальном случае оказываются равными:

$$C_з = \frac{8}{3} \pi \epsilon_0 \epsilon_r L; \quad (5-39a)$$

$$\bar{R}_к = \frac{2}{3} R_{к0}, \quad (5-39b)$$

а минимальная постоянная времени затвора с учетом (5-37) имеет вид:

$$\tau_{з.мин} = C_з \bar{R}_к = \frac{64}{9} \frac{\epsilon_0 \epsilon_r L^2}{a^2}. \quad (5-40)$$

А л к а т р о н характерен кольцевой структурой. Представим себе, что унитрон (см. рис. 5-13, а) вращается вокруг оси y , проходящей через сток. Тогда последний превратится в диск, расположенный в центре конструкции, а затвор и исток превратятся в концентрические кольца, окружающие сток. Такой кольцевой вариант униполярного транзистора (рис. 5-19), предложенный в 1960 г., назван алкатроном [99]. Помимо кольцевой структуры особенностью алкатрона является наличие четвертого электрода — *престриктора*, который существенно улучшает многие параметры прибора.

Главным достоинством алкатрона по сравнению с текстрином является значительно большая рассеиваемая мощность. Это объясняется тем, что в режиме насыщения сечение канала вблизи стока стягивается не в точку, а в линию — окружность довольно большой длины. Кроме того, в алкатроне улучшены условия теплоотвода технологическим обеспечением массивного престриктора. Однако основное назначение престриктора — не теплоотвод, а предварительное сужение (престрикция) канала. Для этого на престриктор, являющийся p -слоем, задается постоянное отрицательное смещение и канал сужается за счет соответствующего расширения нижнего p - n перехода. Таким образом, добавление престриктора позволяет избежать трудностей, связанных с получением очень тонкого канала.

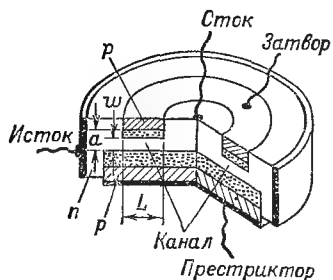


Рис. 5-19. Упрощенная конструкция алкатрона.

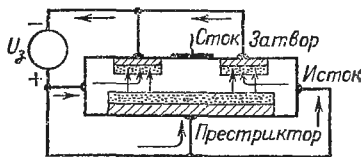


Рис. 5-20. Процесс заряда емкости затвора в алкатроне.

Анализ алкатрона практически совпадает с анализом унитрона, так как длина канала L много меньше радиуса загиба B , и потому специфика кольцевой структуры проявляется слабо. Из формулы (5-28), заменяя величину Z на $2\pi B$, получаем:

$$R_{к0} = \frac{\rho L}{2\pi a B}. \quad (5-41)$$

Здесь величина a отсчитывается от границы престрикторного перехода, и, следовательно, является функцией напряжения $E_{пр}$. Крутизна алкатрона выражается формулой (5-30).

В отношении быстродействия можно заметить следующее. В алкатроне емкость затвора заряжается не только через продольное сопротивление канала, как в унитроне, но и по второй цепи, состоящей из емкости престриктора и p о n е р е ч н о г о сопротивления канала (рис. 5-20). В реальных случаях емкость престриктора больше емкости затвора, а поперечное сопротивление канала меньше продольного. Тогда при прочих равных условиях постоянная времени затвора оказывается у алкатрона меньше, чем у двух других типов полевых транзисторов (см. 2-е изд. данной книги).

Особенности реальных приборов. До сих пор мы рассматривали только процессы в канале, пренебрегая сопротивлениями участков исток — канал $R_{и}$ и канал — сток $R_{с}$. Между тем эти сопротивления могут иметь значительную величину (до 100 Ом и более) и оказывают влияние на параметры приборов. Такое влияние проще всего оценить, присоединяя $R_{и}$ и $R_{с}$ к идеальному транзистору в качестве в н е ш н и х элементов (рис. 5-21).

Как видим, $R_{и}$ играет роль сопротивления обратной связи подобно сопротивлению $R_{к}$ в ламповом каскаде ([62], § 5-8). Поэтому по аналогии с известным в ламповой технике соотношением можно

записать для реальной крутизны транзистора:

$$S = \frac{S^*}{1 + S^* R_{и}}, \quad (5-42)$$

где S^* — крутизна, определенная без учета $R_{и}$ [см. (5-30) и др.].

Постоянная времени затвора тоже изменится, так как, во-первых, заряд емкости перехода будет происходить через суммарное сопротивление $\bar{R}_k + R_{и}$ и, во-вторых, во входной цепи действует обратная связь. Начальная скорость заряда уменьшится в отношении $\bar{R}_k / (\bar{R}_k + R_{и})$, а установившееся значение U_3 — в отношении S/S^* .

Отсюда следует:

$$\tau_3 = \tau_3^* \frac{1 + R_{и}/\bar{R}_k}{1 + S^* R_{и}}, \quad (5-43)$$

где τ_3^* — постоянная времени, определенная в предыдущих разделах [см. (5-32) и др.].

Очевидно, что сопротивление истока следует по возможности уменьшать.

Что касается сопротивления R_c , то оно оказывает некоторое влияние только на крутизну; однако в области насыщения это влияние несущественно. Тем не менее сопротивление стока тоже стараются уменьшить, чтобы снизить рассеяние мощности в приборе и падение напряжения в нем.

Ток через затвор считается обычно равным тепловому току через запертый переход. На самом деле при работе в режиме насыщения ($U_c > U_{c,н}$) ток затвора может иметь более сложную структуру.

Действительно, в «горловине» канала вблизи стока электрическое поле не является плоскопараллельным. Здесь электроны, а также дырки (которые всегда есть в n -канале) движутся по криволинейным траекториям; пучок дырок расширяется в направлении от стока к истоку. Следовательно, дырки имеют составляющие скорости, перпендикулярные оси канала, и частично проходят в область p - n перехода. Здесь они ускоряются и попадают на затвор. Таким образом, положительное приращение потенциала U_3 может вызвать приращение тока I_3 в направлении из затвора. Это означает отрицательное сопротивление затвора, которое в ряде случаев и наблюдается.

Важной особенностью унитронов является их обратимость: исток и сток можно поменять местами, сохранив работоспособность прибора. Однако полная симметрия нормальных и обратных характеристик не всегда имеет место из-за конструктивно-технологических различий областей стока и истока.

Наличие унитронов n - и p -типов со взаимно обратными рабочими полярностями открывает, как и в случае биполярных транзисторов, полезные схемотехнические возможности, хотя унитроны p -типа уступают унитронам n -типа по некоторым электрическим параметрам (см. сноску ¹ на с. 284).

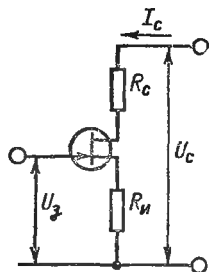


Рис. 5-21. Эквивалентная схема униполярного транзистора с учетом сопротивлений истока и стока.

Унитроны, как и всякий полупроводниковый прибор, подвержены влиянию температуры: от нее зависят ток затвора, ток стока и крутизна. Функция $I_s(T)$ — экспоненциальная (см. § 2-6), так что при температуре 100—125 °C ток I_s может быть в сотни и тысячи раз больше, чем при комнатной температуре.

Функции $I_c(T)$ и $S(T)$ обусловлены зависимостями $\mu(T)$ и $\Delta\phi_0(T)$. Роль зависимости $\mu(T)$ очевидна из выражения (5-28), где $\rho \sim \mu^{-1}$. Учитывая (1-32), приходим к выводу, что $R_{к0} \sim T^c$; следовательно, ток I_c и крутизна S у м е н ь ш а ю т с я с ростом температуры. Это имеет место практически во всех реальных приборах. Однако следует иметь в виду, что в исходных формулах (5-24) и (5-25) для ширины перехода l использовано выражение (2-12) вместо более точного (2-11), в которое входит равновесная высота барьера $\Delta\phi_0$. Последняя согласно (2-5) уменьшается с ростом температуры; вместе с нею уменьшается ширина перехода l , а следовательно, толщина канала a растет. Такая зависимость, как ясно из (5-28), «противодействует» зависимости $\rho(T)$. Анализ показывает [96, 100], что при малых токах можно обеспечить п о с т о я н - с т в о т о к а в д и а п а з о н е т е м п е р а т у р. Условие $dI_c/dT = 0$ выполняется при следующем соотношении:

$$U_{30} - U_3 = 2I_c/S \approx 0,65 \text{ В.} \quad (5-44)$$

В последние годы к униполярным транзисторам проявляется повышенный интерес в связи с их малыми собственными шумами и высоким входным сопротивлением.

Малый уровень шумов обусловлен, с одной стороны, отсутствием инжекции и связанных с ней флуктуаций (см. § 4-10), а с другой стороны, идеальной «изоляцией» канала от поверхности полупроводника благодаря наличию слоя пространственного заряда p - n перехода затвора.

В результате собственные шумы унитронов даже на низких частотах носят в основном тепловой характер, т. е. близки к шумам в весьма малом сопротивлении канала $R_{к0}$. В этом отношении унитроны пока не имеют конкурентов среди полупроводниковых приборов.

Например, на частоте 10 Гц напряжение шумов в полосе 1 Гц может составлять всего 10 нВ при крутизне 4,5 мА/В.

Что касается входного сопротивления, то в кремниевых унитронах, где обратное сопротивление p - n перехода определяется током термогенерации, оно достигает при комнатной температуре 10^{10} — 10^{11} Ом, т. е. превышает входное сопротивление обычных электронных ламп.

Подробное описание и анализ унитронов можно найти в работах [97, 100].

Эквивалентная схема. Аналогия между полевыми транзисторами и электронными лампами объясняет сходство их эквивалентных схем (рис. 5-22).

Строго говоря, канал транзистора и переход затвора нужно было бы представить в виде линий RC с распределенными параметрами, но это, как всегда, затрудняет последующее использование схемы для расчетов. Поэтому на рис. 5-22 и канал, и затвор представлены сосредоточенными параметрами. Канал представлен дифференциальным сопротивлением r_c и межэлектродной емкостью $C_{си}$, которая определяется геометрией и материалом прибора. Затвор представлен сопротивлением r_z и емкостью C_z . Сопротивлением r_z часто пренебрегают в связи с его большой величиной; поэтому на рис. 5-22 оно показано пунктиром. Емкость затвора заряжается через усредненное сопротивление $\bar{R}_K$, которое обуславливает конечную постоянную времени τ_z . Усилительные свойства транзистора отражены генератором тока $\dot{S}U_z$.

Сопротивления истока и стока (R_n и R_c) показаны условно, так как их влияние (см. выше с. 291) автоматически учитывается при измерении параметров S , τ_z и r_c со стороны внешних зажимов.

Типичными значениями параметров кремниевых унитронов (в режиме $U_{си} = 10$ В, $U_{зи} = 0$) являются: $S = 0,3 \div 3$ мА/В, $r_z = 10^{10}$ Ом, $r_c = 0,1 \div 1$ МОм, $\bar{R}_K = 75 \div 200$ Ом, $C_z = 3 \div 10$ пФ, $C_{зи} = 0,5$ пФ, $C_{сз} = 0,5$ пФ, $C_{си} = 0,3 \div 1$ пФ. Фактор шума минимален при сопротивлении источника сигнала $R_r = 5 \div 10$ МОм и может составлять $F_{ш} < 1$ дБ на частоте 100 Гц и менее.

Инерционность транзистора отражается не только наличием емкостей, но и тем, что крутизна S является комплексной или операторной величиной, которую приближенно можно выразить следующим образом:

$$S(s) \approx \frac{S}{1 + s\tau_s}, \quad (5-45)$$

где $\tau_s \approx \tau_z$.

Типичными значениями τ_s и граничной частоты f_s являются соответственно 0,5—2 нс и 100—300 МГц.

5-5. ПОЛЕВЫЕ ТРАНЗИСТОРЫ С ИЗОЛИРОВАННЫМ ЗАТВОРОМ

Структура и классификация. Спецификой унитронов является максимальная проводимость канала при нулевом смещении на затворе. С ростом смещения (по модулю) проводимость канала

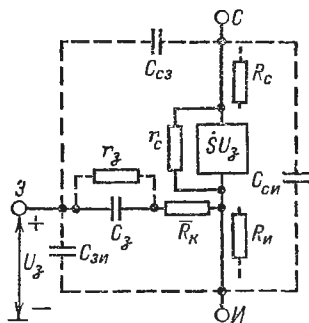


Рис. 5-22. Эквивалентная схема униполярного полевого транзистора для переменных составляющих токов и напряжений.

уменьшается вплоть до полной отсечки. Смещение может иметь только одну полярность, соответствующую отсутствию инжекции через переход.

У полевых транзисторов с изолированным затвором последний представляет собой металлический слой, изолированный от полупроводника тонкой диэлектрической пленкой. Наличие диэлектрика снимает ограничение на полярность смещения: она может быть как положительной, так и отрицательной, причем в обоих случаях ток затвора отсутствует. Структура таких транзисторов (металл — диэлектрик — полупроводник), как уже отмечалось, лежит в основе их названия: *МДП транзисторы*. В том весьма распространенном случае, когда диэлектриком является окисел (двуокись кремния), их называют *МОП транзисторами* (по-английски *MOS*).

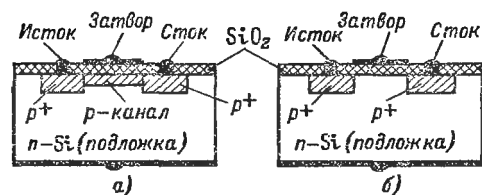


Рис. 5-23. Структуры МДП транзисторов с собственным (а) и индуцированным (б) каналами.

В случае канала *p*-типа положительный потенциал U_z «отталкивает» дырки из канала (*режим обеднения*), а отрицательный — «притягивает» их (*режим обогащения*). Соответственно проводимость канала либо уменьшается, либо увеличивается по сравнению с ее значением при нулевом смещении.

Вторая структура (рис. 5-23, б) характерна отсутствием структурно выраженного канала. Поэтому при нулевом смещении на затворе проводимость между истоком и стоком практически отсутствует: исток и сток образуют с подложкой встречно-включенные *p-n* переходы. Тем более не может быть существенной проводимости между истоком и стоком при *п о л о ж и т е л ь н о й* полярности смещения, когда к поверхности полупроводника притягиваются дополнительные электроны. Однако при достаточно большом отрицательном смещении, когда приповерхностный слой сильно обогащается притянутыми дырками, между истоком и стоком образуется *индуцированный* (наведенный полем) канал, по которому может протекать ток. Значит, транзисторы с индуцированным каналом работают только в режиме обогащения. В настоящее время этот тип транзисторов имеет наибольшее распространение.

Физические процессы. У каждой разновидности МДП транзисторов имеются свои, иногда существенные особенности, в зависимости от типа канала (встроенный или индуцированный), типа проводимости канала (*p* или *n*), а также от типа диэлектрика, полу-

проводника и металла. Мы подробно проанализируем работу кремниевого МОП транзистора с индуцированным p -каналом, а относительно других разновидностей ограничимся краткими замечаниями.

МОП транзисторы с p -каналом характерны отрицательными полярностями рабочих напряжений, поэтому ниже напряжения и потенциалы записываются в виде модулей, чтобы избежать неоднократного употребления знака минус.

Прежде всего отметим, что в равновесном состоянии, т. е. в отсутствие приложенных напряжений, приповерхностный слой окисленного кремния (как n -, так и p -типа) всегда обогащен электронами. Соответственно зоны в этом слое искривлены вниз, а на границе с окислом поверхность имеет отрицательный потенциал φ_{so} (рис. 5-24, а). Причины обогащения приповерхностного слоя электронами следующие: контактная разность потенциалов между металлом затвора и кремнием; поверхностные уровни донорного типа; положительные ионы в диэлектрике¹. Первые две причины знакомы по § 2-4. Что касается положительных ионов, то они, как показывает опыт, появляются в процессе окисления кремния и последующей фотолитографической обработки окисла (с. 267). Наиболее распространены ионы натрия и водорода, причем, как правило, они сосредоточены вблизи границы окисла с кремнием. Разумеется, наличие обогащенного электронами слоя затрудняет образование дырочного канала в рассматриваемом типе тран-

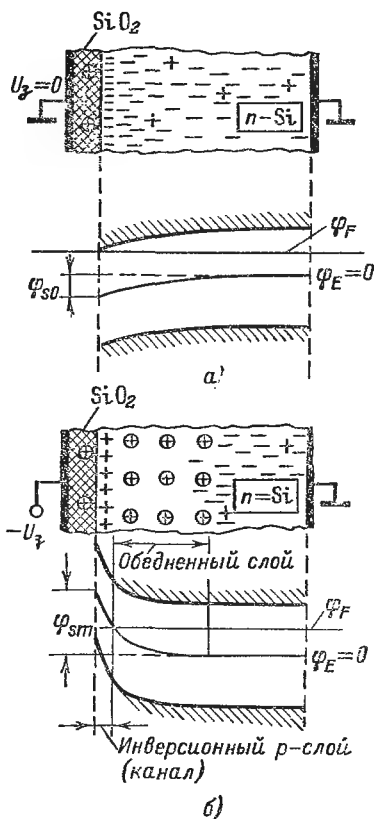


Рис. 5-24. Распределение носителей и зонные диаграммы в МОП транзисторе с индуцированным p -каналом.

а — равновесное состояние; б — при отрицательном смещении на затворе.

¹ Вообще говоря, контактная разность потенциалов может иметь и положительную полярность; тогда этот фактор противодействует двум другим, т. е. способствует уменьшению отрицательного приповерхностного заряда или даже изменению его знака. Однако для наиболее распространенной комбинации «алюминий — кремний» контактная разность потенциалов отрицательна (независимо от типа проводимости кремния, см. табл. 2-1) и, следовательно, способствует обогащению приповерхностного слоя электронами.

зисторов¹. По мере роста отрицательного смещения U_3 электроны отталкиваются от поверхности. При этом энергетические зоны сначала спрямляются, а затем искривляются вверх, т. е. поверхностный потенциал делается положительным. Однако до тех пор, пока поверхностная область сохраняет электронный тип проводимости (точнее, пока сток и исток образуют с этой областью выпрямляющие контакты), проводимость рабочей цепи остается крайне малой и протекание заметного тока по-прежнему невозможно.

Существует некоторое *пороговое напряжение* $U_3 = U_0$, по превышении которого энергетические зоны искривляются настолько сильно, что вблизи поверхности образуется инверсионный дырочный слой (см. § 2-4). Именно этот слой играет роль индуцированного канала (рис. 5-24, б). Принято считать [100, 101], что пороговое напряжение соответствует поверхностному потенциалу $\varphi_{sm} = 2\varphi_F$, где φ_F есть «расстояние» между уровнем Ферми и уровнем электростатического потенциала в глубине полупроводника (на рис. 5-24 в этой области принято $\varphi_E = 0$).

Дальнейший рост напряжения U_3 слабо влияет на величину φ_s , поскольку изменения последней всего на несколько φ_T уже вызывают изменения концентрации дырок в десятки раз [см. (1-176), где φ_E следует заменить на φ_s].

Наряду с образованием дырочного канала под затвором образуется также обедненный слой, в котором положительный заряд обусловлен обнаженными ионами доноров (рис. 5-24, б). Образование обедненного слоя вызвано отталкиванием основных носителей подложки — электронов — от поверхности.

Инверсионный слой значительно тоньше обедненного. Толщина последнего составляет сотни ангстрем и более, а толщина индуцированного канала — всего 10—20 ангстрем.

Рассмотрим динамику образования канала и прохождения тока в нем.

Если в начальном состоянии все электроды МДП транзистора находились при нулевом потенциале, а затем на затвор была подана ступенька напряжения $U_3 > U_0$, то в первый момент поле будет иметь примерно такую конфигурацию, как показано на рис. 5-25, а. Под действием этого поля электроны перемещаются в сторону от поверхности обнажая ионы доноров и образуя обедненный слой, а дырки движутся к поверхности, накапливаются вблизи границы с диэлектриком и образуют канал². Равновесие наступает тогда, когда поле, обусловленное дыр-

¹ В МОП транзисторах с n -каналом наличие равновесного обогащенного электронами слоя равносильно наличию встроенного канала. Для того чтобы получить прибор с индуцированным n -каналом, нужно использовать подложку из весьма низкоомного p -кремния, в котором концентрация дырок превышает концентрацию электронов в приповерхностном обогащенном слое. Однако низкоомный материал ухудшает параметры транзистора (растут емкости, падает пробивное напряжение).

² Из рис. 5-25, а следует, что дырки поступают в канал не только из подложки, но также из полуметаллических r^+ -слоев истока и стока. Именно эти слои и являются главными «поставщиками» дырок, поскольку напряженность «тянущего» поля вблизи них максимальна, а концентрация дырок в них практически неограниченна, тогда как в подложке n -типа дырок мало и они генерируются сравнительно медленно (с постоянной времени τ).

ками и ионами доноров, компенсирует исходное поле в подложке, а также на границах истока и стока. Можно считать, что описанный переходный процесс соответствует заряду емкости затвора. В установившемся состоянии основная часть подложки нейтральна, а поле ограничено узким участком под затвором: силовые линии начинаются на дырках и ионах доноров (а также на ионах в диэлектрике)

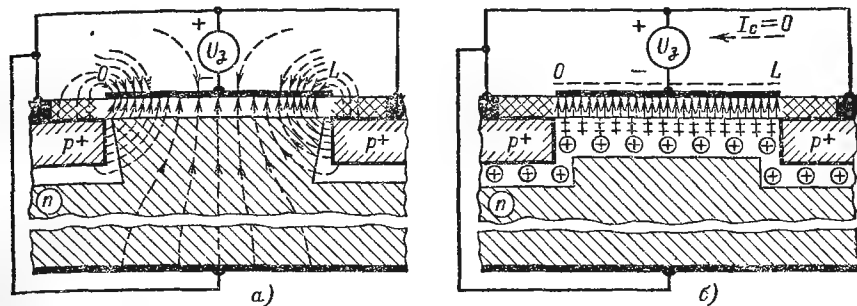


Рис. 5-25. Распределение поля в МДП транзисторе в отсутствие напряжения на стоке.

a — в первый момент; *б* — в установившемся состоянии.

и кончаются на электроде затвора (на рис. 5-25, *б* для простоты все силовые линии показаны начинающимися на границе полупроводник — диэлектрик). Поле в диэлектрике однородное, поскольку потенциал всей поверхности полупроводника одинаков и равен $\Phi_s \approx \Phi_{sm}$.

Если теперь, после образования канала, подать ступенчатое напряжение отрицательной полярности между стоком и истоком (считая, что последний по-

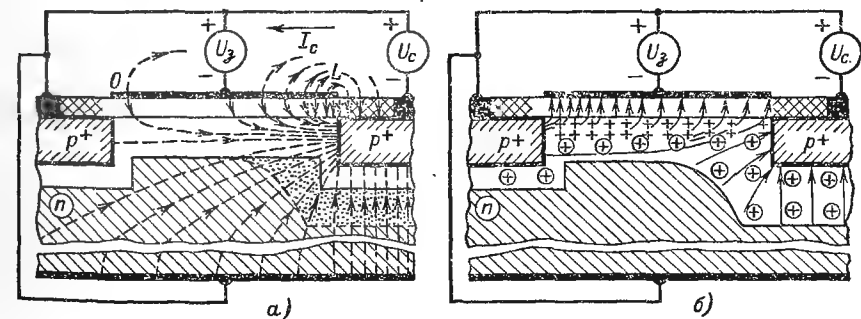


Рис. 5-26. Распределение поля в МДП транзисторе при подаче напряжения на сток.

a — в первый момент (показаны линии поля, обусловленные только напряжением U_c); *б* — в установившемся состоянии.

прежнему соединен с подложкой), то в первый момент появляется дополнительное поле, конфигурация которого показана на рис. 5-26, *a*. Свободные носители начнут двигаться в этом поле, дырки — в сторону стока, электроны — в противоположном направлении. Перемещение электронов в подложке сопровождается обнажением дополнительных ионов доноров (т. е. расширением объемного заряда вокруг стока) и соответствующим нарастанием «противополя», обусловленного этими ионами. Данный процесс (процесс заряда барьерной емкости стока) закончится тогда, когда результирующее поле в подложке снова

сделается равным нулю, а силовые линии стока будут замыкаться уже не на электрод подложки, а на слой дополнительного заряда донорных ионов (на рис. 5-26, а этот слой показан точками, см. также рис. 5-26, б).

Что касается дырок, то их «отсос» из канала приводит к изменению удельного заряда поверхности (имеется в виду заряд на единицу площади затвора). Согласно теореме Гаусса удельный заряд поверхности однозначно связан с напряженностью поля в диэлектрике: $Q_0 = \epsilon_0 \epsilon_{дл} E$. Следовательно, поле в диэлектрике становится **н е о д н о р о д н ы м**: его напряженность убывает от истока к стоку. Такой вывод вполне естествен и основан на том простом факте, что потенциал поверхности под затвором меняется от φ_{sm} (на границе истока) до $U_c + \varphi_{sm}$ (на границе стока).

Если «сложить» поля, показанные на рис. 5-26, а и 5-25, б то результирующее поле в установившемся состоянии (по окончании заряда барьерной емкости

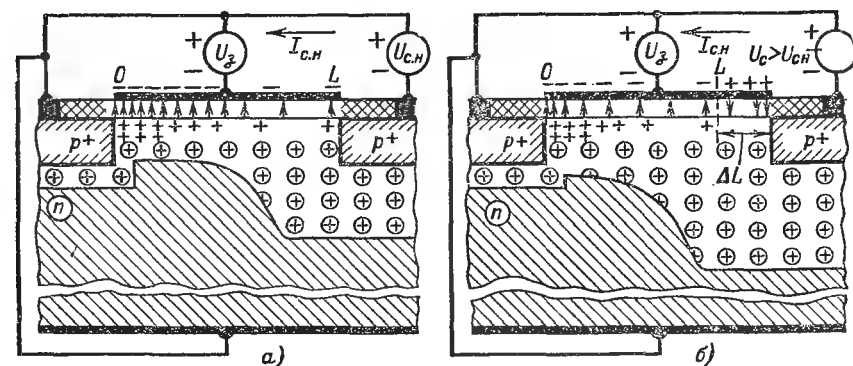


Рис. 5-27. Распределение поля в диэлектрике МДП транзистора.

а — на границе насыщения; б — в режиме насыщения.

стока) будет иметь примерно такую конфигурацию, как показано на рис. 5-26, б. Как видим, поле в диэлектрике действительно неоднородное, а в самом канале появилась **т а н г е н ц и а л ь н а я** составляющая поля, которая обеспечивает движение дырок (ток), а также их пополнение из p^+ -слоя истока¹.

При некотором критическом потенциале стока, который называется **напряжением насыщения** $U_{c,n}$, поле в диэлектрике и удельный заряд в полупроводнике в **б л и з и** **с т о к а** делаются равными нулю (рис. 5-27, а). Это значит, что на границе стока образуется «горловина» дырочного канала или, иными словами, в этом месте обедненный слой доноров «выходит на поверхность». При значениях $U_c > U_{c,n}$ обедненный слой на поверхности расши-

¹ Заметим, что силовые линии между истоком и стоком отсутствуют и, следовательно, движение дырок происходит не под действием поля исток — сток, а под действием поля исток — затвор. Иначе говоря, дырочный канал нельзя уподобить проводящей пластинке, расположенной между истоком и стоком. Роль потенциала U_c состоит в том, что он, во-первых, увеличивает напряженность поля в обедненном слое, окружающем сток (это вызывает «отсос» дырок, расположенных в этом слое близ поверхности, и тем самым инициирует движение дырок), а во-вторых, придает полю между истоком и затвором такую конфигурацию, при которой имеется тангенциальная составляющая, необходимая для движения дырок (в частности, для их «вытягивания» из истока).

руется пропорционально $\sqrt{U_c - U_{c.н}}$, как в обычном $p-n$ переходе, а «горловина» канала, сохраняющая потенциал $U_{c.н}$, соответственно сдвигается в сторону от стока. Вследствие этого происходит некоторое укорочение канала (рис. 5-27, б). Аналогичное явление отмечалось при анализе унитронов (см. рис. 5-15).

Семейство выходных характеристик МДП транзистора показано на рис. 5-28, а. С увеличением напряжения на стоке (при постоянном потенциале U_3) ток I_c сначала нарастает почти линейно (пока дырочный канал вблизи стока слабо деформирован), затем нарастание тока замедляется (поскольку канал вблизи стока су-

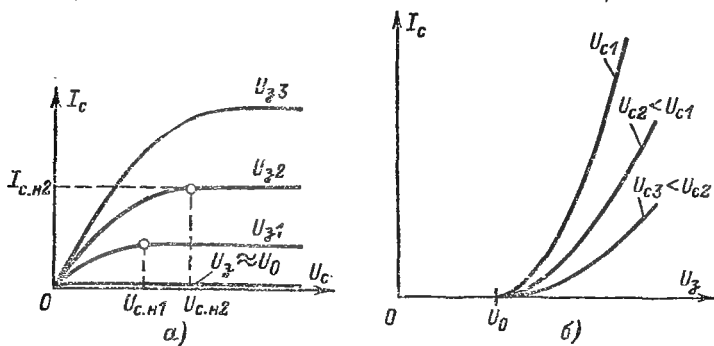


Рис. 5-28. Статические выходные (а) и передаточные (б) характеристики МДП транзистора.

жается и его сопротивление растет) и, наконец, получается участок насыщения, где ток остается почти постоянным (небольшой рост тока обусловлен уменьшением длины канала, см. выше). Если увеличить напряжение U_3 , то канал обогатится дырками, его исходное сопротивление уменьшится и соответствующая кривая $I_c(U_c)$ пройдет выше. Насыщение тока в этом случае наступит позднее, так как исходный удельный заряд дырок стал больше и, чтобы свести его к нулю, требуется большее напряжение $U_{c.н}$.

В целом семейство выходных вольт-амперных характеристик МДП транзистора напоминает семейство выходных характеристик унитрона (см. рис. 5-16, а). Главное отличие состоит в том, что с ростом параметра U_3 у МДП транзистора ток увеличивается (режим обогащения), а у унитронов уменьшается (режим обеднения). Кроме того, на семействе МДП транзистора отсутствует кривая с параметром $U_3 = 0$, поскольку канал индуцируется только при условии $U_3 > U_0$. При значениях $U_3 < U_0$ в цепи стока протекает очень небольшой остаточный ток — ток обратного смещенного стокового перехода.

Что касается передаточных характеристик $I_c(U_3)$, то они у МДП транзистора имеют такой же вид, как у унитрона, но сдвинуты относительно нуля координат на величину порогового напряжения (рис. 5-28, б).

Общий анализ. Задача анализа — найти зависимость тока I_c от напряжений U_a и U_c . При этом исток считается соединенным с подложкой, которая имеет нулевой потенциал.

На рис. 5-29 показано распределение положительных зарядов в канале и обедненном слое МДП транзистора. Поскольку ток обусловлен только подвижными носителями, установим связь между током и зарядом дырок в канале. Учтывая, что ток одинаков в любом сечении, рассмотрим произвольное сечение с координатой x .

Считая ток чисто дрейфовым, запишем его плотность в соответствии с (1-72a) в виде

$$j_c = q\mu_p p(x, y) E_x,$$

где напряженность E_x принята независимой от координаты y (в пределах крайне малой толщины канала h это допустимо).

Поскольку распределение $p(y)$ неизвестно, воспользуемся усредненным значением $p_{cp}(x)$, которое представляет собой

интеграл функции $p(x, y)$ по y (в пределах от 0 до h), поделенный на h . Замена $p(x, y)$ на $p_{cp}(x)$ дает усредненную (по оси y) плотность тока. Умножив ее на площадь Zh , получим выражение для тока в следующем виде:

$$I_c = \mu_p Z Q_{op}(x) E_x. \quad (5-46)$$

Здесь $Q_{op}(x) = qhp_{cp}(x)$ по размерности и физическому смыслу есть удельный заряд дырок, приходящийся на единицу площади в плоскости xz (или, образно говоря, расположенный под этой площадью). Дальнейшая задача состоит в определении заряда Q_{op} .

Рассмотрим баланс удельных зарядов в МДП транзисторе (рис. 5-30) *. В равновесном состоянии ($U_s = 0$) вблизи поверхности имеется обогащенный слой (см. начало предыдущего раздела). Отрицательный заряд обогащенного слоя балансируется положительными зарядами на затворе и на границе Si — SiO₂ (рис. 5-30, а). Первый обусловлен контактной разностью потенциалов, а второй — ионами в диэлектрике и поверхностными состояниями донорного типа. Если подать на затвор отрицательное напряжение $U_s = \Phi_{MS}$, то заряд на затворе делается равным нулю, а в обогащенном слое исчезнет («оттолкнется») составляющая, связанная с контактной разностью потенциалов (рис. 5-30, б).

* Там, где это не вызывает недоразумений, опускаем для краткости слово «удельный».

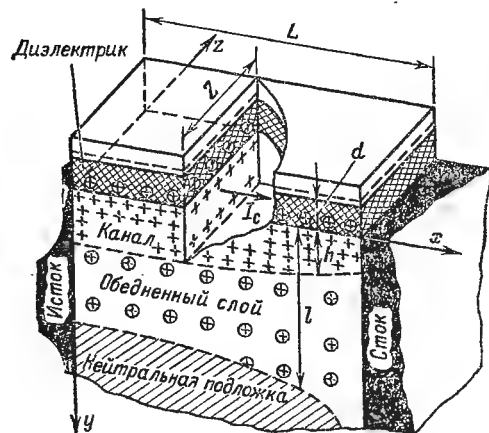


Рис. 5-29. Распределение зарядов в структуре МДП транзистора.

Если увеличить напряжения U_3 (по модулю) до значения U_F (напряжение спрямления зон, см. с. 70), то обогащенный слой полностью ликвидируется, а заряд на затворе становится отрицательным, равным сумме зарядов ионов и поверхностных состояний (рис. 5-30, в). Дальнейшее увеличение U_3 сопровождается образованием обедненного слоя, увеличением положительного заряда доноров и соответствующим увеличением отрицательного заряда на затворе; при этом поверхностный потенциал ϕ_s делается положительным и возрастает до значения $\phi_{sm} \approx 2\phi_F$, когда образуется канал.

В момент образования канала напряжение на затворе равно пороговому напряжению U_0 (рис. 5-30, г). При значениях $U_3 > U_0$ заряд в дырочном канале растет вместе с напряжением U_3 ; соответственно растет отрицательный заряд на затворе. Поскольку при наличии канала потенциал ϕ_s меняется мало (см. предыдущий раздел, с. 296), то заряд обедненного слоя, а вместе с ним и соответствующая составляющая заряда на затворе тоже остаются практически неизменными (рис. 5-30, д).

Итак, из проведенного рассмотрения следует, что удельный заряд дырок в канале можно найти из соотношения

$$|Q_{03}| = Q_{0p} + Q_{0d} + Q_{0п.с}, \quad (5-47)$$

где $Q_{0п.с}$ — суммарный заряд поверхностных состояний и ионов (его обычно не разделяют на составляющие и считают заданной величиной, полученной из измерений). Запишем выражения для зарядов Q_{03} и Q_{0d} , входящих в (5-47).

Обозначая потенциал поверхности полупроводника в точке x через U_x , приходим к выводу, что напряженность поля в диэлектрике будет $[(U_3 - \phi_{ms}) - U_x]/d^*$. Тогда согласно теореме Гаусса удельный заряд затвора при данном значении x запишется следующим образом:

$$Q_{03} = \epsilon_0 \epsilon_d E = \frac{\epsilon_0 \epsilon_d}{d} (U_3 - \phi_{ms} - U_x), \quad (5-48)$$

где ϵ_d — относительная диэлектрическая проницаемость диэлектрика.

* Обозначение U_x вместо ϕ_s для поверхностного потенциала призвано подчеркнуть, что в данном случае речь идет не о квазиравновесном состоянии

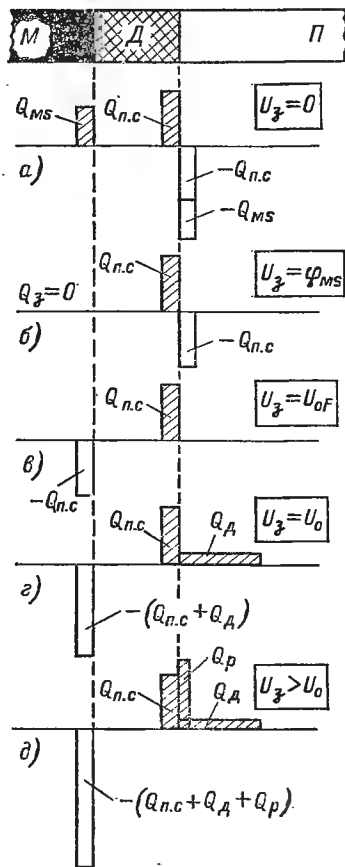


Рис. 5-30. Баланс зарядов в структуре МДП транзистора при характерных напряжениях на затворе.

Что касается величины $Q_{0д}$, то для ее определения нужно было бы проинтегрировать по координате y функцию $qN_d^*(y)$, где N_d^* — концентрация «обнаженных» донорных ионов. Такая задача, как уже отмечалось в § 1-12, весьма сложна в математическом отношении. Поэтому воспользуемся аппроксимацией, которая использовалась при анализе ступенчатого p - n перехода в гл. 2, а именно, положим $N_d^* = N_d = \text{const}$, считая, что подвижные носители в обедненном слое полностью отсутствуют. Тогда по аналогии с (1-90б) получаем для толщины обедненного слоя:

$$l = \sqrt{\frac{2\epsilon_0\epsilon_n U_x}{qN_d}}.$$

Искомую величину $Q_{0д}$ найдем, умножая толщину l на объемную плотность заряда qN_d :

$$Q_{0д} = qN_d l = a \sqrt{U_x}, \quad (5-49)$$

где

$$a = \sqrt{2q\epsilon_0\epsilon_n N_d}; \quad (5-50)$$

ϵ_n — относительная диэлектрическая проницаемость полупроводника.

Подставляя в (5-46) заряд $Q_{0р}$, выраженный с помощью (5-47) через три других заряда, используя соотношения (5-48) и (5-49) и заменяя E_x на $-dU_x/dx$, получаем следующее дифференциальное уравнение¹:

$$I_c = \mu Z C_0 \left(U_3 - U_{0F} - U_x - \frac{a}{C_0} \sqrt{U_x} \right) \frac{dU_x}{dx}, \quad (5-51)$$

где

$$C_0 = \frac{\epsilon_0 \epsilon_d}{d} \quad (5-52)$$

— удельная емкость затвора относительно поверхности полупроводника;

$$U_{0F} = \frac{Q_{0п.с}}{C_0} + \varphi_{MS} \quad (5-53)$$

— напряжение спрямления зон (с. 70), т. е. то напряжение на затворе, которое меняет потенциал поверхности на величину φ_{s0} , обеспечивая значение $\varphi_s \approx 0$ (см. рис. 5-24, а).

(когда нет тока и напряжение подается только на затвор), а о неравновесном состоянии, когда $U_c \neq 0$ и протекает ток.

Использование разности $U_3 - \varphi_{MS}$ вместо U_3 объясняется тем, что нулевой заряд на затворе получается при условии $U_3 = \varphi_{MS}$ (см. рис. 5-30, б); следовательно отрицательные заряды, характерные для рабочего режима транзистора, соответствуют не полному потенциалу U_3 , а разности $U_3 - \varphi_{MS}$.

¹ Знак минус в правой части (5-51) опущен, так как под U_x понимается модуль отрицательного потенциала поверхности.

Например, если диэлектрическим является двуокись кремния ($\epsilon_d = 3,5$), то при толщине $d = 0,1$ мкм, контактной разности потенциалов $\varphi_{ms} = 0,3$ В и плотности поверхностных состояний $2 \cdot 10^{11}$ см⁻² (т. е. $Q_{оп.с} \approx 3 \cdot 10^{-8}$ Кл/см²) получаем: $C_0 \approx 3 \cdot 10^{-8}$ Ф/см² (30 000 пФ/см²) и $U_{0F} \approx 1,3$ В.

Умножив обе части уравнения (5-51) на dx и проинтегрировав левую часть в пределах от 0 до L , а правую — от φ_{sm} (см. рис. 5-26, б) до $\varphi_{sm} + U_c$, найдем искомую вольт-амперную характеристику:

$$I_c = b \left\{ (U_s - U_{0F} - \varphi_{sm}) U_c - \frac{1}{2} U_c^2 - \frac{2}{3} \frac{a}{C_0} [(\varphi_{sm} + U_c)^{3/2} - \varphi_{sm}^{3/2}] \right\}. \quad (5-54)$$

Параметр b , определяющий «масштаб» характеристики, называется *удельной крутизной* и имеет значение¹

$$b = \frac{\mu Z C_0}{L} = \frac{\epsilon_0 \epsilon_d \mu Z}{L d}. \quad (5-55)$$

При типичных параметрах $\mu = 200$ см²/В·с, $C_0 = 3 \cdot 10^{-8}$ Ф/см² (см. выше) и $Z/L = 10$ получаем $b = 60$ мкА/В².

Выражение (5-54) неудобно для инженерных расчетов из-за наличия членов в степени 3/2, однако из этого выражения можно сделать полезные выводы.

Прежде всего определим *пороговое напряжение* МДП транзистора (см. с. 296), при котором образуется дырочный канал. Для этого сначала найдем из (5-54) проводимость канала dI_c/dU_c ; затем подставим $U_c = 0$, что соответствует отсутствию тока, т. е. одинаковой толщине канала вдоль оси x ; наконец, положим $dI_c/dU_c = 0$ — условие, означающее исчезновение проводящего канала. Из этого условия нетрудно получить пороговое напряжение на затворе; запишем его в следующем виде:

$$U_0 = U_{0F} + U_{0B}. \quad (5-56)$$

Здесь U_{0B} — составляющая порогового напряжения, соответствующая значению $\varphi_s = \varphi_{sm}$, т. е. обеспечивающая искривление зон «вверх», достаточное для образования канала:

$$U_{0B} = \frac{a}{C_0} \sqrt{\varphi_{sm}} + \varphi_{sm}. \quad (5-57)$$

Назовем величину U_{0B} в отличие от U_{0F} *напряжением изгиба зон*.

Величина U_{0B} определяет минимальное значение порогового напряжения, когда отсутствуют такие «привходящие» факторы, как поверхностные состояния в полупроводнике и ионы

¹ По ряду причин, обусловленных влиянием границы раздела, приповерхностная подвижность носителей в формуле (5-55) обычно в 2—3 раза меньше объемной.

В диэлектрике, т. е. когда $U_{0F} = 0$. Например, для кремниевого МДП транзистора при концентрации доноров $N_d = 10^{15} \div 10^{16} \text{ см}^{-3}$ коэффициент a согласно (5-50) составляет $(2-6) \cdot 10^{-8} \text{ В}^{1/2} \text{ Ф/см}^2$, а потенциал Ферми φ_F согласно (1-21в) — около 0,3 В. Тогда при $C_0 = 3 \cdot 10^{-8} \text{ Ф/см}^2$ и $\varphi_{sm} = 2\varphi_F$ (см. с. 296) напряжение U_{0B} лежит в пределах 1—2 В.

Полное пороговое напряжение U_0 , определяемое формулой (5-56), у современных МДП транзисторов обычно составляет 2,5—4В.

Из выражения (5-54) следует, что функция $I_c(U_c)$ имеет максимум при некотором напряжении $U_{c,n}$, которое можно найти путем дифференцирования. Правее максимума ток согласно (5-54) должен был бы уменьшаться. На самом деле он сохраняет максимальное значение, так как при условии $U_c > U_{c,n}$ качественно меняется механизм работы МДП транзистора (см. предыдущий раздел) и выражение (5-54) теряет силу. Следовательно, выражение (5-54) описывает только начальные (*крутые*) участки вольт-амперных характеристик (см. рис. 5-28, а), где ток зависит от обоих напряжений U_3 и U_c . В режиме насыщения (на *пологих* участках) ток практически зависит только от напряжения U_3 , а влияние напряжения U_c нуждается в специальном анализе (см. ниже).

В заключение отметим, что в n -канальных транзисторах с подложкой p -типа начальный обогащенный слой тоже электронный (см. сноску на с. 295); поэтому напряжение плоских зон U_{0F} будет по-прежнему отрицательным¹, тогда как напряжение изгиба зон U_{0B} положительно. Если $|U_{0F}| > U_{0B}$, то $U_0 < 0$, т. е. транзистор будет иметь канал даже при $U_3 = 0$ и, следовательно, будет аналогичен МДП транзистору со встроенным каналом, хотя технологически такой канал не предусмотрен (об этой трудности уже упоминалось в сноске¹ на с. 296). При условии $|U_{0F}| < U_{0B}$ обеспечивается индуцированный канал, причем пороговое напряжение U_0 будет меньше, чем у p -канальных транзисторов. Однако воспроизводимость величины U_0 обычно хуже из-за разброса напряжения U_{0F} , обусловленного трудно контролируруемыми факторами.

Характеристики и параметры в 1-м приближении. Рассмотрим самую простую аппроксимацию выражения (5-54), которой удобно пользоваться при инженерных расчетах. Заменим сумму $U_{0F} + \varphi_{sm}$ на полное пороговое напряжение U_0 , а членами с коэффициентом a/C_0 пренебрежем. Тогда

$$I_c = b \left[(U_3 - U_0) U_c - \frac{1}{2} U_c^2 \right]. \quad (5-58)$$

Такая аппроксимация всегда действительна при малых сток-вых напряжениях ($U_c < \varphi_{sm}$). В этом легко убедиться, разлагая функцию $(\varphi_{sm} + U_c)^{3/2}$ в ряд Маклорена с точностью до двух первых членов, подставляя их в выражение (5-54) и учитывая (5-57) и (5-56). При значениях $U_c > \varphi_{sm}$ аппроксимация (5-58) действительна лишь в случае достаточно малых значений a/C_0 , т. е. в случае

¹ Напомним, что применительно к p -канальным транзисторам мы пользуемся модулями зарядов и напряжений. На самом деле напряжение U_0 и его компоненты в формуле (5-56) имеют отрицательные значения.

тонких слоев диэлектрика [см. (5-52)] и высокочастотных подложек с малой концентрацией примеси [см. (5-50)]. Расчеты показывают, что аппроксимация (5-58) дает погрешность не более 30%, если $U_3 - U_0 > 0,5$ В и $a/C_0 < 1$ В^{1/2}. Для кремниевых МОП транзисторов последнее ограничение обычно соответствует концентрации $N_d < 2 \cdot 10^{15}$ см⁻³, т. е. удельному сопротивлению подложки $\rho_n > 4$ Ом·см.

С физической точки зрения аппроксимация (5-58) действительна тогда, когда заряд обедненного слоя не играет существенной роли в работе прибора: например, если толщина этого слоя не меняется вдоль канала (случай $U_c < \psi_{sm}$) или если удельный заряд слоя много меньше удельного заряда затвора (случай малых значений a/C_0).

Из выражения (5-58) легко найти напряжение насыщения, полагая $dI_c/dU_c = 0$:

$$U_{c,n} = U_3 - U_0. \quad (5-59)$$

Как уже отмечалось, значение тока $I_c(U_{c,n})$ сохраняется при всех значениях $U_c > U_{c,n}$. Поэтому, подставляя (5-59) в (5-58), получаем вольт-амперную характеристику МДП транзистора в режиме насыщения:

$$I_c = \frac{1}{2} b (U_3 - U_0)^2. \quad (5-60)$$

В усилительной технике МДП транзисторы всегда используются в режиме насыщения, поскольку ему свойственны наименьшие нелинейные искажения и оптимальные значения дифференциальных параметров: крутизны $S = \partial I_c / \partial U_3$, внутреннего сопротивления $R_i = \partial U_c / \partial I_c$ и собственного коэффициента усиления $\bar{\mu} = \partial U_c / \partial U_3$. Эти параметры связаны между собой «ламповым» соотношением

$$\bar{\mu} = S R_i \quad (5-61)$$

(черточка над μ призвана отличить данный параметр от подвижности).

Крутизна в режиме насыщения легко определяется из выражения (5-60):

$$S = b (U_3 - U_0). \quad (5-62)$$

Как видим, крутизна линейно зависит от *эффективного напряжения* на затворе $U_3 - U_0$, а при заданном эффективном напряжении пропорциональна параметру b . Название последнего (удельная крутизна) обусловлено тем, что при $U_3 - U_0 = 1$ В величина b численно равна крутизне в режиме насыщения. Вообще же параметр b является производной $\partial^2 I_c / \partial U_3 \partial U_c$, т. е. характеризует *кривизну* функции $I_c(U_3; U_c)$.

Крутизна МДП транзистора однозначно связана с током. Эту связь легко установить с помощью выражений (5-60) и (5-62):

$$S = \sqrt{2bI_c}. \quad (5-63)$$

Из выражения (5-62) с учетом (5-55) ясно, что крутизна МДП транзистора возрастает с уменьшением толщины диэлектрика d , а также с увеличением ширины канала Z . Однако оба эти фактора одновременно способствуют увеличению емкости между затвором и каналом, которая описывается выражением

$$C_s = C_0(ZL) = \frac{\epsilon_0 \epsilon ZL}{d}. \quad (5-64)$$

Поэтому такой глобальный параметр всякого усилительного прибора, как *добротность* [62], определяемая отношением $S/2,2 C$, не зависит от величин d и Z . Поделив (5-62) на (5-64), получаем добротность МДП транзистора в следующем виде:

$$D = \frac{\mu}{2,2L^2} (U_s - U_0). \quad (5-65)$$

Отсюда следует, что основным средством повышения добротности является уменьшение длины канала L .

В качестве примера положим $b = 100$ мкА/В², $C_0 = 3 \cdot 10^{-8}$ Ф/см², $L = 10^{-3}$ см (10 мкм) и $Z = 2 \cdot 10^{-2}$ см (200 мкм). Тогда емкость затвора $C_s \approx 0,6$ пФ, а крутизна и добротность (при эффективном напряжении $U_s - U_0 = 3$ В) соответственно равны: $S \approx 0,3$ мА/В и $D \approx 0,5$ нс⁻¹.

В отличие от крутизны внутреннее сопротивление и коэффициент усиления МДП транзистора не могут быть определены из выражения (5-60), поскольку оно не содержит напряжения U_c . Согласно (5-60) характеристики в режиме насыщения должны быть горизонтальными ($R_i = \infty$, $\bar{\mu} = \infty$).

В разделе «Физические процессы» отмечалось, что конечное значение R_i в режиме насыщения обусловлено зависимостью ширины стокового перехода от напряжения U_c — явлением, не учтенным в формуле (5-60). Учет этого явления осложняется неоднородностью электрического поля в стоковом переходе. Если такой неоднородностью пренебречь и принять для функции $l(U_c)$ выражение (2-12), свойственное ступенчатому p - n переходу, то, дифференцируя по U_c выражение (5-60) (в котором параметр b является функцией длины канала L) и полагая $dL/dU_c = -dl/dU_c$, можно привести внутреннее сопротивление к следующему виду¹:

$$R_i = \frac{L}{l} \frac{U_c}{I_c} = \frac{L \sqrt{U_c}}{I_c} \sqrt{\frac{2qN_d}{\epsilon_0 \epsilon_n}}. \quad (5-66)$$

Тогда из соотношения (5-61) с учетом (5-63), (5-55) и (5-64) получается выражение для коэффициента усиления:

$$\bar{\mu} = 2 \sqrt{\frac{U_c}{I_c}} \sqrt{\frac{qN_d \mu C_s}{\epsilon_0 \epsilon_n}}. \quad (5-67)$$

Как видим, параметры R_i и $\bar{\mu}$ увеличиваются с ростом рабочего напряжения U_c и с уменьшением рабочего тока (т. е. при малых значениях $U_s - U_0$).

¹ Более подробно и строго дифференциальные параметры МДП транзисторов в режиме насыщения рассмотрены в работах [102, 103].

Разумеется, в области достаточно больших напряжений U_c наступают предпробойные явления, а затем и пробой, сопровождающийся резким возрастанием тока I_c и столь же резким уменьшением сопротивления R_i . Заметим, что пробой может иметь место не только в стоковом p - n переходе, но также и в диэлектрике (между стоком и затвором).

В импульсных схемах МДП транзистор работает в качестве ключа и основной интерес представляют две крайние рабочие точки, соответствующие запертому и максимально открытому состоянию ключа (рис. 5-31).

Запертое состояние (точка 1) характеризуется условием $U_3 < U_0$. При этом в цепи стока протекает лишь некоторый остаточный ток, обусловленный утечками по поверхности, а также обратным током p - n перехода стока (если подложка находится под нулевым или положительным потенциалом). В качественных МОП транзисторах этот ток не превышает нескольких наноампер.

Максимально открытое состояние ключа (точка 2) характеризуется большими значениями эффективного напряжения U_3 — U_0 и расположением рабочей точки на крутом участке соответствующей вольт-амперной кривой. Обычно в открытом состоянии ток I_c оказывается заданным внешней цепью, а интерес представляет остаточное напряжение U_{c0} . Это напряжение легко определить из формулы (5-58), однако в общем виде выражение получается громоздким. На практике остаточное напряжение настолько мало, что в формуле (5-58) можно пренебречь членом $\frac{1}{2} U_c^2$. Тогда

$$U_{c0} = \frac{I_c}{b(U_3 - U_0)}. \quad (5-68)$$

Выражение (5-68) действительно при условии $U_3 - U_0 > (2-3) U_{c0}$.

Еще одним параметром, важным для характеристики открытого ключа, является сопротивление на начальном участке кривой $I_c(U_c)$. Дифференцируя (5-58) по U_c , получаем дифференциальное выходное сопротивление транзистора в ненасыщенном режиме:

$$R_i = \frac{1}{b(U_3 - U_0 - U_c)}. \quad (5-69a)$$

Если положить $U_c \ll U_3 - U_0$, получим искомое сопротивление открытого ключа:

$$R_0 = \frac{1}{b(U_3 - U_0)}. \quad (5-69b)$$

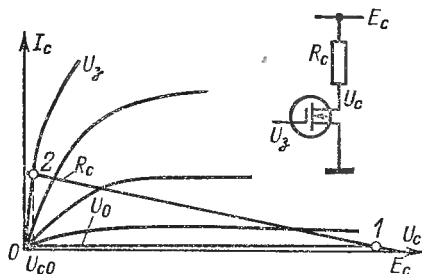


Рис. 5-31. Рабочие точки МДП транзистора в ключевом режиме.

Это же выражение можно получить непосредственно из формулы (5-68). Поскольку при малых значениях U_c сечение дырочного слоя и концентрация дырок в нем почти не зависят от x , величину R_0 обычно называют *сопротивлением канала*.

Из (5-68) и (5-69б) очевидно, что остаточное напряжение на МДП транзисторном ключе имеет чисто «омическое» происхождение:

$$U_{c0} = I_c R_0.$$

Это обстоятельство является важным преимуществом МДП транзисторов, так как в случае биполярных транзисторов даже при нулевом токе имеется вполне конечное остаточное напряжение [см. (15-10в)].

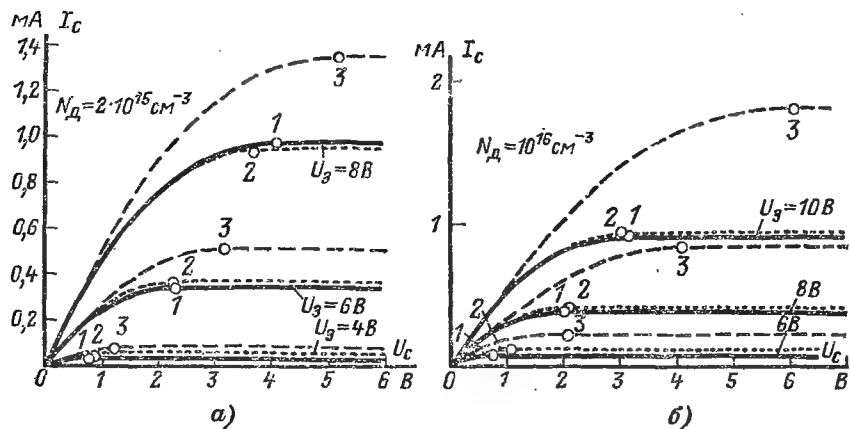


Рис. 5-32. Вольт-амперные характеристики МДП транзистора, построенные по (5-54) (сплошные линии), (5-58) (крупный пунктир) и (5-70) (мелкий пунктир). 1—3 — точки насыщения соответствующих кривых: а — случай сравнительно высокоомной подложки; б — случай сравнительно низкоомной подложки.

Характеристики и параметры во 2-м приближении. Аппроксимация (5-58) и связанные с ней выражения вполне пригодны для приближенных расчетов и качественных оценок. Однако напряжение насыщения, а значит, и ток насыщения определяются при этом со значительной погрешностью, особенно в случае низкоомных подложек. Эта погрешность может достигать 50% (рис. 5-32). Еще большая погрешность получается тогда, когда потенциал подложки отличен от нуля, что часто имеет место в интегральных схемах. Поэтому иногда приходится использовать более точные аппроксимации, которые рассматриваются ниже.

Пусть потенциал подложки по-прежнему равен нулю, но желательно отразить влияние членов в степени $3/2$, входящих в выражение (5-54), сохранив при этом удобную форму выражения (5-58). Наиболее простой путь решения такой задачи состоит в следующем.

Вместо 1-го приближения (5-58) примем 2-е:

$$I_c = b \left[(U_s - U_0) U_c - \frac{1}{2} (1 + \eta) U_c^2 \right], \quad (5-70)$$

в котором поправочный коэффициент η нужно выбрать так, чтобы напряжения насыщения $U_{c.n.}$, определенные из (5-70) и (5-54), были достаточно близки друг к другу. Поскольку напряжение $U_{c.n.}$ получается из условия $dI_c/dU_c = 0$, продифференцируем правые части (5-70) и (5-54) и приравняем эти производные. Тогда коэффициент η после некоторых преобразований можно записать в виде

$$\eta = \frac{a}{C_0 V \varphi_{sm}} \frac{\sqrt{m+1}-1}{m},$$

где $m = U_c/2\varphi_{sm}$. Поскольку выражение (5-70) действительно лишь для ненасыщенного режима (т. е. при $U_c < U_{c.n.}$), не имеет смысла рассматривать сколь угодно большие значения m . Также не имеет смысла рассматривать значения $m < 1$, так как при этом действительно 1-е приближение (5-58) (см. с. 304). Разумным диапазоном можно считать $m = 2 \div 15$. Тогда, как показывают расчеты, коэффициент η можно принять не зависящим от напряжения U_c и равным:

$$\eta = \frac{1}{3} \frac{a/C_0}{V \varphi_{sm}}. \quad (5-71)$$

Используя параметры, характерные для кремниевых МОП транзисторов, получаем: $\eta = 0,8 \cdot 10^{-3} d \sqrt{N_d}$. Например, если $d = 10^{-5}$ см и $N_d = 10^{16}$ см⁻³ (т. е. $\rho_n \approx 0,8$ Ом·см), то $\eta \approx 0,8$. Из рис. 5-32 видно, что аппроксимация (5-70) обеспечивает высокую точность.

Дифференцируя (5-70) по U_c и полагая $dI_c/dU_c = 0$, находим напряжение насыщения:

$$U_{c.n.} = \frac{U_s - U_0}{1 + \eta}. \quad (5-72)$$

Оно оказывается меньше, чем при расчете по (5-59). Подставляя (5-72) в (5-70), получаем уточненную характеристику МДП транзистора в режиме насыщения [ср. с (5-60)]:

$$I_c = \frac{1}{2} b' (U_s - U_0)^2, \quad (5-73)$$

где

$$b' = \frac{b}{1 + \eta} \quad (5-74)$$

— эквивалентная удельная крутизна, несколько меньшая реальной (5-55).

Естественно, что крутизна S и добротность D , определяемые величиной b' , будут во 2-м приближении тоже в $1 + \eta$ раз меньше,

чем в 1-м:

$$S = \frac{b}{1+\eta} (U_3 - U_0); \quad (5-75)$$

$$D = \frac{\mu}{2,2(1+\eta)L^2} (U_3 - U_0). \quad (5-76)$$

Для ключевого режима, характерного малыми напряжениями U_c , остаются в силе выражения (5-68) и (5-69), полученные из 1-го приближения.

Для тех (сравнительно редких) случаев, когда желательно описать вольт-амперные характеристики в целом, отразив и насыщенный и ненасыщенный режим в [104] предложена эмпирическая аппроксимация:

$$I_c = b (U_3 - U_0)^2 \frac{\text{th}(U_c/U_{c.n}) [1 - 0,5 \text{th}(U_c/U_{c.n})]}{1 - \frac{2}{\pi L} \sqrt{\frac{2\epsilon_0 e n}{q N_d}} \sqrt{U_c} \text{arcth}(U_c/U_{c.n})}.$$

Эта формула, в которой $U_{c.n} = U_3 - U_0$, в предельных режимах $U_c \ll U_{c.n}$ и $U_c \gg U_{c.n}$ дает результаты, близкие соответственно к (5-58) и (5-60); в то же время она отражает конечный наклон кривых в режиме насыщения (и даже факт роста сопротивления R_I с увеличением напряжения U_c , см. (5-66)).

Влияние потенциала подложки. Пусть подложка имеет относительно истока положительный потенциал U_n . Тогда падение напряжения на обедненном слое при координате x будет уже не U_x , как считалось в формуле (5-49), а $U_x + U_n$. Соответственно и в формуле (5-51) нужно заменить U_x на $U_x + U_n$. Эта поправка, как легко убедиться, приводит к росту составляющей порогового напряжения U_{0B} : вместо (5-57) получаем выражение

$$U_{0B} = \frac{a}{C_0} \sqrt{\varphi_{sm} + U_n} + \varphi_{sm}, \quad (5-77)$$

а вместо (5-56) — выражение

$$U_0 = U_{0F} + \frac{a}{C_0} \sqrt{\varphi_{sm} + U_n} + \varphi_{sm}. \quad (5-78)$$

Характеристика (5-54) остается в силе с соответствующими поправками:

$$I_c = b \left\{ (U_3 - U_{0F} - \varphi_{sm}) U_c - \frac{1}{2} U_c^2 - \frac{2}{3} \frac{a}{C_0} [(\varphi_{sm} + U_n + U_c)^{3/2} - (\varphi_{sm} + U_n)^{3/2}] \right\}. \quad (5-79)$$

Что касается аппроксимаций, то оба приближения (5-58) и (5-70) остаются в силе, если для порогового напряжения использовать значение (5-78). Кроме того, в выражении (5-71) следует заменить φ_{sm} на $\varphi_{sm} + U_n$:

$$\eta' = \frac{1}{3} \frac{a/C_0}{\sqrt{\varphi_{sm} + U_n}}. \quad (5-80)$$

Как видим, с ростом потенциала U_n коэффициент η' уменьшается.

Очевидно, что в принципе потенциал U_n может управлять током I_c подобно потенциалу U_z . Крутизну по подложке $S_n = \partial I_c / \partial U_n$ нетрудно найти из характеристики (5-79) или из характеристик (5-58) и (5-70) [с учетом (5-78) и (5-80)]. Однако выражения для S_n получаются громоздкими и ненаглядными. Поэтому имеет смысл аппроксимировать характеристику (5-79) линейной функцией от U_n .

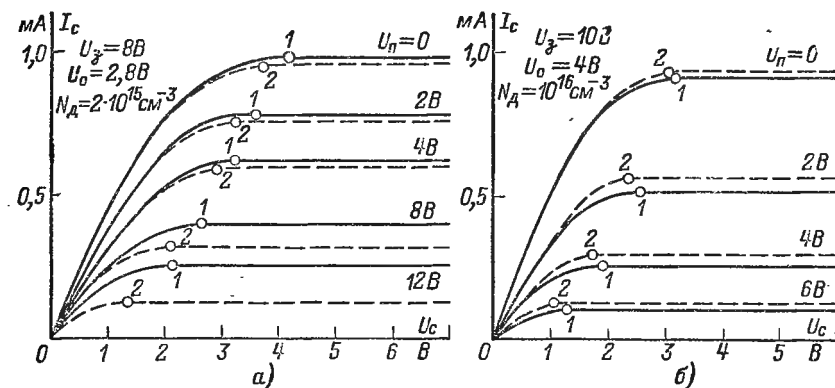


Рис. 5-33. Вольт-амперные характеристики МДП транзистора, построенные по (5-79) (сплошные линии) и по (5-81) (пунктир).

1 и 2 — точки насыщения соответствующих кривых: а — случай сравнительно высокоомной подложки; б — случай сравнительно низкоомной подложки.

Наиболее простая аппроксимация имеет следующий вид¹ (рис. 5-33):

$$I_c = b \left[\left(U_z - U_0 - \frac{2}{3} \eta U_n \right) U_c - \frac{1}{2} (1 + \eta) U_c^2 \right], \quad (5-81)$$

где коэффициент η имеет значение (5-71), не зависящее от U_n , а пороговое напряжение U_0 определяется формулами (5-56) и (5-57), т. е. без учета потенциала подложки. Из (5-81) обычным

¹ В [105, 106] предлагались другие аппроксимации, например

$$I_c = b \left\{ [U_z - (1 + \eta) U_0 - \eta U_n] U_c - \frac{1}{2} (1 + \eta) U_c^2 \right\},$$

где $\eta = 0,5 \frac{a/C_0}{\sqrt{U_0}}$. Такая аппроксимация представляется менее обоснованной, хотя бы потому, что в ней η оказывается функцией U_0 . Автор работы [105] предложил для режима насыщения дальнейшее уточнение приведенной формулы:

$$I_c = \frac{1}{2} \frac{b}{1 + \eta} \left(1 + 5\lambda^2 \frac{U_c - U_{c,n}}{U_0} \right) [U_z - (1 + \eta) U_0 - \eta U_n]^2,$$

где $\lambda = \frac{1}{\eta} \frac{e_n}{e_d} \frac{d}{L}$ [этот коэффициент отражает роль внутреннего сопротивления R_i и получен из тех же соображений, что и (5-66)].

способом находим приращение насыщения:

$$U_{c.u} = \frac{1}{1+\eta} \left(U_s - U_0 - \frac{2}{3} \eta U_n \right). \quad (5-82)$$

Подставляя (5-82) в (5-81), получаем обобщение формулы (5-73) для режима насыщения:

$$I_c = \frac{1}{2} b' \left(U_s - U_0 - \frac{2}{3} \eta U_n \right)^2. \quad (5-83)$$

Наконец, дифференцируя (5-83) по U_n , определяем крутизну по подложке:

$$S_n = -\frac{2}{3} \eta b' \left(U_s - U_0 - \frac{2}{3} \eta U_n \right). \quad (5-84)$$

Знак минус говорит о том, что ток I_c возрастает с уменьшением положительного потенциала U_n , т. е. при отрицательных приращениях ΔU_n .

Крутизна по затвору согласно (5-83) равна:

$$S = b' \left(U_s - U_0 - \frac{2}{3} \eta U_n \right). \quad (5-85)$$

Сравнивая (5-84) и (5-85), видим, что соотношение крутизн зависит от коэффициента η , т. е. в конечном счете от толщины диэлектрика и удельного сопротивления подложки. У обычных кремниевых МОП транзисторов $|S_n| < S$. Если соединить затвор с подложкой, то результирующая крутизна будет равна сумме $S + |S_n|$.

Следует иметь в виду, что входное сопротивление по подложке определяется обратными токами p - n переходов и, следовательно, несравненно меньше входного сопротивления по затвору.

Эквивалентная схема. Эквивалентная схема для переменных составляющих должна отражать влияние подложки (независимо от наличия потенциала U_n). Одна из таких схем показана на рис. 5-34. Несмотря на внешнюю симметрию, обе половины схемы имеют разные значения параметров. Например, как отмечалось в конце предыдущего раздела, крутизна S обычно больше, чем S_n , а сопротивления $R_{зи}$ и $R_{зс}$ намного больше, чем $R_{пи}$ и $R_{пс}$. Межэлектродные емкости $C_{зи}$ и $C_{зс}$ в значительной мере обусловлены тем, что металлический электрод затвора в реальных структурах расположен не точно между слоями истока и стока (как на рис. 5-23 и др.), а частично — краями — над ними. Такая структурная осо-

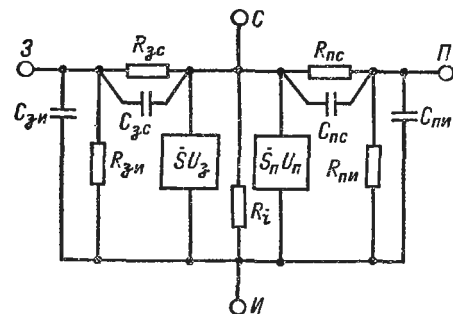


Рис. 5-34. Эквивалентная схема МДП транзистора для малых переменных составляющих.

бенность называется *перекрытием затвора*. Значения указанных емкостей зависят от степени перекрытия, т. е. от того, насколько длина металлической обкладки затвора превышает длину канала (обычно на 2—3 мкм). Емкости $C_{зи}$ и $C_{зс}$, составляющие 0,5—1 пФ на 1 мм ширины канала, в несколько раз меньше емкостей p - n переходов стока и особенно истока $C_{пи}$ и $C_{пс}$.

Заметим, что емкость $C_{зс}$ является элементом обратной связи между выходом и входом транзистора; поэтому уменьшению этой емкости всегда уделяется большое внимание при разработке приборов. В последние годы удалось разработать новые варианты технологии (так называемые методы *самосовмещенных затворов*), которые позволили на порядок, до долей пикофарады, уменьшить паразитную емкость $C_{зс}$ [100, 107] путем ликвидации перекрытия затвора.

В режиме насыщения, характерном для линейных усилительных схем, крутизна современных МОП транзисторов обычно составляет 1—2 мА/В при токе около 1 мА, хотя, увеличивая отношение Z/L , можно согласно (5-62) и (5-55) добиться значительно больших значений¹. Внутреннее сопротивление R_i в режиме насыщения обычно лежит в пределах 50—100 кОм (при том же токе), а коэффициент усиления μ — соответственно в пределах 100—200.

Малый коэффициент усиления — один из главных недостатков МДП транзисторов по сравнению с биполярными (у последних он достигает нескольких тысяч, см. стр. 335—336).

Одной из важнейших особенностей МДП транзисторов является их огромное входное сопротивление (на рис. 5-34 отражено резисторами $R_{зи}$, $R_{зс}$). Обычно оно составляет не менее 10^{14} Ом, а при использовании специальных мер [108] доходит до 10^{17} Ом и выше. Значительно меньшее значение (10^9 — 10^{10} Ом) имеют сопротивления $R_{пс}$ и $R_{пи}$, т. е. обратные сопротивления p - n переходов истока и стока. Примерно такое же значение имеет сопротивление R_i в режиме отсечки.

Особо нужно сказать о пробое затвора. Слой диэлектрика под затвором настолько тонок (около 0,1 мкм), а его сопротивление настолько велико, что пробой легко наступает от накопления статического электричества, например, при простом прикосновении к свободному («плавающему») затвору.

Переходные и частотные характеристики МДП транзисторов обусловлены перезарядкой межэлектродных емкостей через внешние резисторы, а также перезарядкой емкости затвор — канал через сопротивление канала. Последний процесс накладывает принци-

¹ Обычно в простейшей, *полосковой* конструкции отношение Z/L трудно сделать более 100 из-за различных «перекосов» при фотолитографии. Однако, используя *змейковую* или *гребенчатую* структуру (см. рис. 4-35), равносильную параллельному соединению многих транзисторов, удастся разработать приборы на токи в несколько ампер и мощности в десятки ватт. Крутизна таких приборов составляет сотни миллиампер на вольт и более.

пиальное ограничение на быстродействие транзистора. Строго говоря, канал МДП транзистора следует рассматривать как распределенную RC -систему. Однако удобнее в первом приближении охарактеризовать инерционность этой системы постоянной времени, т. е. уподобить канал простейшей RC -цепочке. В качестве элементов такой цепочки можно принять емкость затвора (5-64) и сопротивление канала (5-69). Тогда постоянная времени канала (она же *постоянная времени крутизны*) будет иметь вид:

$$\tau_S = R_0 C_3 = \frac{L^2}{\mu (U_3 - U_0)}. \quad (5-86)$$

Как видим, это по существу обратная величина добротности (5-65). Для улучшения переходных и частотных свойств МДП транзистора нужно в первую очередь уменьшать длину канала, а также увеличивать приповерхностную подвижность. При обычной длине канала 5—10 мкм граничная частота крутизны ($1/2 \pi \tau_S$) лежит в пределах 100—300 МГц. Однако при ультракоротких (доли микрона) каналах удастся обеспечить граничные частоты до 10—20 ГГц. На эквивалентной схеме параметры C_3 и R_0 не показаны, так как их влияние отражено операторным или комплексным характером крутизны S . В типичных случаях емкость затвора составляет десятые доли пикофарда (иногда 1—2 пФ), а сопротивление канала — сотни ом.

Переходя к температурным зависимостям, можно отметить, что МДП транзисторы, как и унитроны (см. стр. 292), характерны наличием так называемого критического тока $I_{c,кр}$, при котором температурная чувствительность тока dI_c/dT равна нулю. Значение критического тока (или соответствующего ему напряжения $U_3 - U_0$) получается путем дифференцирования выражений (5-60) или (5-73), с учетом зависимостей $b(T)$ и $U_0(T)$. Показано [100], что зависимость $b(T)$ связана с функцией $\mu(T)$, а зависимость $U_0(T)$ — с функцией $\varphi_{sm}(T)$. В результате анализа, с учетом выражений (1-32) и (1-13), оказывается, что условие $dI_c/dT = 0$ соответствует эффективному напряжению на затворе:

$$U_3 - U_0 = (0,8 - 2,4) \text{ В}$$

(минимальное значение соответствует концентрации примеси в подложке 10^{18} см^{-3} , максимальная — концентрации 10^{15} см^{-3}). Обычно ток $I_{c,кр}$ в 5—10 раз меньше номинального тока¹. В диапазоне $I_c > I_{c,кр}$ (в частности, в номинальном режиме) $dI_c/dT > 0$, а в области микрорежима ($I_c < I_{c,кр}$), наоборот, $dI_c/dT < 0$.

На практике температурную стабильность МДП транзисторов характеризуют чаще не приращением тока ΔI_c , а эквивалентным приращением потенциала затвора ΔU_3 , которое обеспечивает постоянство тока I_c при изменениях температуры. Оба приращения связаны очевидным соотношением

¹ Номинальным считают ток, соответствующий напряжению $U_3 - U_0 = U_0$; с учетом (5-60) $I_{ном} = \frac{1}{2} b U_0^2$.

$\Delta I_c = S \Delta U_s$. Отсюда величина температурной чувствительности ϵ_U (т. е. производной dU_s/dT) связана с температурной чувствительностью ϵ_I (т. е. производной dI_c/dT) соотношением

$$\epsilon_U = \frac{\epsilon_I}{S}. \quad (5-87)$$

Для токов, близких к критическому, характерны значения $\epsilon_U = \pm 0,5 \text{ мВ/}^\circ\text{C}$, для «сверхкритических» токов (в частности, номинальных) $\epsilon_U = + (8 - 10) \text{ мВ/}^\circ\text{C}$, а для «субкритических» (в 10—20 раз меньших критического) $\epsilon_U = - (4 - 6) \text{ мВ/}^\circ\text{C}$.

Проводя аналогичный анализ применительно к крутизне, нетрудно убедиться, что можно получить $\epsilon_S \approx 0$, но при меньшем токе, чем $I_{c, \text{кр}}$. В обоих случаях осложняющим обстоятельством является то, что малая или нулевая ϵ получается ценою снижения временной стабильности рабочей точки, так как критические напряжения $(U_s - U_0)_{\text{кр}}$ оказываются малыми по сравнению с напряжением U_0 и дрейф последнего сильно сказывается на разности $U_s - U_0$. Более подробные сведения о температурном и временном дрейфе можно найти в [109, 100].

В заключение отметим тенденцию к уменьшению порогового напряжения МДП транзисторов с тем, чтобы приблизить их напряжения питания к напряжениям питания биполярных транзисторов в интегральных схемах (3—5 В). За последнее время на этом пути достигнуты большие успехи. Одним из новых методов является замена алюминия в качестве электрода затвора на материалы, обеспечивающие меньшую контактную разность потенциалов с кремнием [см. (5-53)]. Например, используется молибден или поликристаллический кремний, который при достаточном легировании имеет сравнительно небольшое удельное сопротивление. Такие приборы называют МДП транзисторами с молибденовым или кремниевым затвором [110]. Их пороговые напряжения доходят до 0,5—1 В.

Схемы включения. МДП транзистор, как и биполярный (и как любой другой активный трехполюсник), может включаться тремя разными способами. По аналогии со схемами ОЭ, ОК и ОБ у биполярных транзисторов (см. рис. 4-4) у МДП транзисторов различают схемы с *общим истоком* (ОИ), *общим стоком* (ОС) и *общим затвором* (ОЗ) (рис. 5-35). Первая, наиболее распространенная, подробно исследована выше. Для остальных двух ограничимся краткими замечаниями.

С х е м а ОС лежит в основе *истоковых повторителей*, которые, как и другие *повторители*, характерны повышенным входным и пониженным выходным сопротивлениями (по сравнению с основной схемой ОИ), а также коэффициентом передачи напряжения, близким к единице. В схеме ОС выходное сопротивление (со стороны истока) нетрудно найти, задавая на выходе некоторое напряжение ΔU и определяя соответствующий ток $\Delta I \approx \Delta I_u$. Поскольку потенциалы U_s и U_c неизменны, получаем:

$$|\Delta U_{\text{зи}}| = |\Delta U_{\text{сн}}| = \Delta U.$$

Подставляя эти значения в формулу

$$\Delta I_n = \Delta I_c = S \Delta U_{зи} + \frac{1}{R_i} \Delta U_{си}. \quad (5-88)$$

и используя соотношение (5-61), запишем выходное сопротивление $\Delta U / \Delta I$ в следующем виде:

$$R_{\text{вых ОС}} = \frac{R_i}{\mu + 1}. \quad (5-89a)$$

При условии $\mu \gg 1$, которое всегда выполняется,

$$R_{\text{вых ОС}} \approx \frac{1}{S}. \quad (5-89б)$$

Входное сопротивление в схеме ОС (без нагрузки в цепи истока), разумеется, равно R_g .

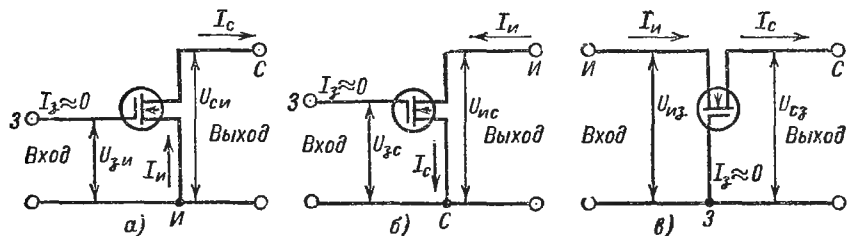


Рис. 5-35. Схемы включения МДП транзистора.

а — с общим истоком; б — с общим стоком; в — с общим затвором.

С х е м а ОЗ, как и схема ОБ, обычно не имеет самостоятельного значения, но находит применение в качестве компонента сложных усилительных схем, например так называемых *каскадов* (см. § 10-4). Она характерна низким входным сопротивлением, которое (как легко видеть из сравнения схем ОЗ и ОС) равно выходному сопротивлению в схеме ОС. Поэтому, учитывая (5-89), запишем:

$$R_{\text{вх ОЗ}} = \frac{R_i}{\mu + 1} \approx \frac{1}{S}. \quad (5-90)$$

Что касается выходного сопротивления, то в схеме ОЗ оно такое же, как и в схеме ОИ, т. е. равно R_i . Это легко показать с помощью (5-88), учитывая, что при определении $R_{\text{вых}}$ в схеме ОЗ потенциалы U_n и U_a остаются постоянными, т. е. $\Delta U_{зи} = 0$.

Глава шестая

СТАТИЧЕСКИЙ РЕЖИМ УСИЛИТЕЛЬНОГО КАСКАДА

6-1. ВЫБОР РАБОЧЕЙ ТОЧКИ

В режиме покоя усилительного каскада, когда нет сигнала, нужно правильно выбрать рабочую точку транзистора, т. е. совокупность тока покоя I_k и напряжения покоя U_k . На семействах характеристик ОБ и ОЭ, показанных на рис. 6-1 вместе с соответствующими схемами каскадов, рабочие точки обозначены буквой А. Через эту точку, как известно, проходят обе линии нагрузки: статическая R_k и динамическая $R_k \parallel R_n$ (см. [111, рис. 2-23]).

Рабочая точка выбирается, исходя из заданных максимальных амплитуд выходного напряжения (U_m) и связанного с ним тока [$I_m = U_m / (R_k \parallel R_n)$], а именно

$$U_{kA} > U_m; \quad I_{kA} > I_m. \quad (6-1)$$

Для линейных усилительных каскадов неравенства (6-1) должны выполняться достаточно сильно. Помимо этого рабочая точка должна удовлетворять условиям

$$U_{kA} + U_m < U_{k, \text{доп}}; \quad U_{kA} I_{kA} < P_{k, \text{доп}}, \quad (6-2)$$

т. е. должна лежать левее вертикали $U_{k, \text{доп}}$ и ниже гиперболы $P_{k, \text{доп}}$, где $U_{k, \text{доп}}$ и $P_{k, \text{доп}}$ — допустимые напряжение и мощность (рис. 6-1).

В случае малого сигнала (десятые доли вольта и менее) выполнение соотношений (6-1) не встречает затруднений, так что рабочая точка может выбираться из условий максимального коэффициента β , максимальной граничной частоты f_a , достаточно малой потребляемой мощности и т. п. Чаще всего используется режим, рекомендованный в справочниках.

Определив желательные координаты рабочей точки (U_{kA} , I_{kA}), нужно обеспечить их в реальной схеме, выбрав соответствующие напряжения источников питания и смещения, а также номиналы режимных резисторов. Такие расчеты можно выполнить аналитически.

Поскольку наибольшее распространение имеют каскады, в которых транзистор работает по схеме ОЭ, возьмем за основу эквивалентную схему на рис. 4-22. Введем в нее для общности источники э. д. с. и внешние сопротивления в цепи всех трех электро-

дов¹. Тогда получается обобщенная схема каскада (рис. 6-2, а), действительная для постоянных составляющих токов и напряжений при любом включении транзистора (ОБ, ОЭ, ОК). Здесь собственное сопротивление базы r_6 входит в сопротивление R_6 , а падение напряжения на эмиттерном переходе отражено генерато-

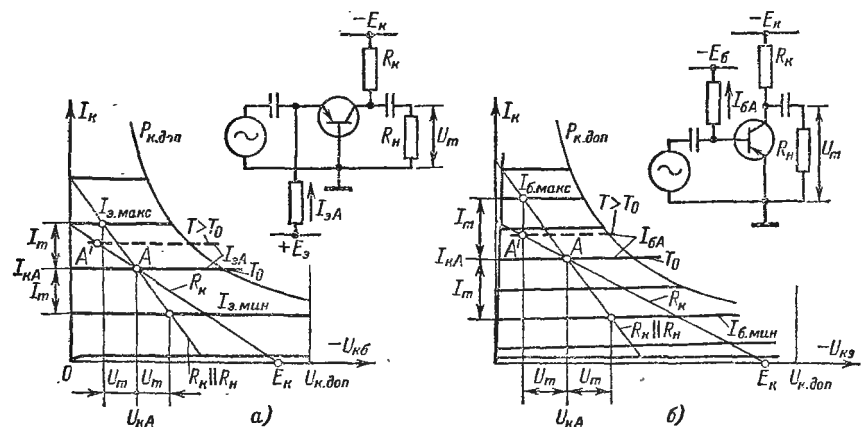


Рис. 6-1. Расположение рабочей точки A на коллекторных характеристиках.
а — в каскаде ОБ; б — в каскаде ОЭ.

ром э. д. с. $U_{эб}$. Обобщенная схема позволяет найти токи и напряжения в каскаде.

Определим ток базы, воспользовавшись принципом суперпозиции. Положим сначала $\beta = 0$; $I_{к0} = 0$ (холостой ход генераторов тока), а затем $E_э = E_6 = E_к = U_{эб}' = 0$ (короткое замыкание генераторов э. д. с.) и найдем соответствующие составляющие базового тока. Ток I_6 будет суммой этих составляющих, из которых первая обусловлена действием э. д. с. $E_э + E_6 - U_{эб}'$ во входном контуре, а вторая — ответвлением тока $I_к$ в цепь базы. Результат можно записать в следующем виде:

$$I_6 = \frac{E_{эб} - U_{эб}'}{R_{эб}} - \gamma_B I_к. \quad (6-3a)$$

Здесь $E_{эб} = E_э + E_6$ — суммарная внешняя э. д. с. в контуре эмиттер — база; $R_{эб} = R_э + R_Б$ — суммарное сопротивление в том же контуре:

$$\gamma_B = \frac{R_э}{R_э + R_Б} \quad (6-3б)$$

— коэффициент токораспределения, показывающий, какая часть

¹ Влияние сопротивления $r_к^*$ при расчетах режима не очень существенно, но вносит осложнения в анализ. Поэтому роль этого сопротивления будет оценена позднее, в гл. 13.

тока I_K ответвляется в базу. Подставляя (6-2а) в выражение для коллекторного тока (4-72) и решая относительно I_K , получаем:

$$I_K = \frac{\beta (E_{эб} - U_{эб}) / R_{эб} + I_{к0}^*}{1 + \beta \gamma_B}. \quad (6-3в)$$

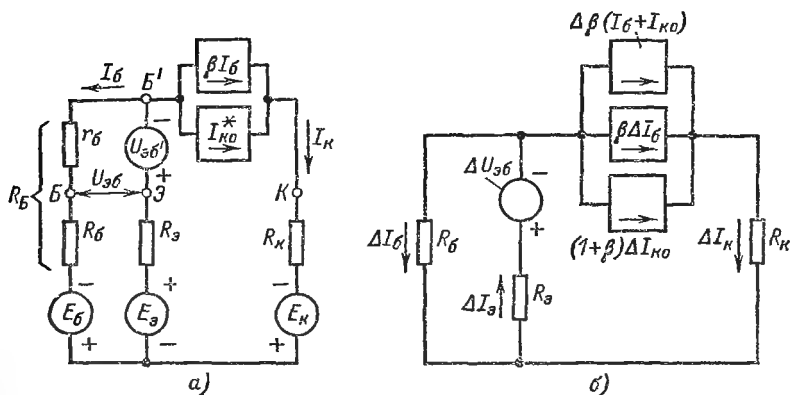


Рис. 6-2. Обобщенные эквивалентные схемы каскада для постоянных составляющих (а) и для приращений (дрейфа) постоянных составляющих (б).

Если в исходное выражение (6-3а) подставить $U_{эб}' = U_{эб} - I_б r_б$ и решить получающееся уравнение относительно $I_б$, то предыдущие выражения переходят в следующие:

$$I_б = \frac{E_{эб} - U_{эб}}{R_{эб}} - \gamma_б I_K; \quad (6-4а)$$

$$\gamma_б = \frac{R_э}{R_э + R_б}; \quad (6-4б)$$

$$I_K = \frac{\beta (E_{эб} - U_{эб}) / R_{эб} + I_{к0}^*}{(1 + \beta \gamma_б)}. \quad (6-4в)$$

Здесь $R_{эб} = R_э + R_б$, а напряжение $U_{эб}$ находят по входным характеристикам при токе $I_{эА} \approx I_{кА}^*$.

Как видим, при условии $r_к = \infty$, которое мы приняли, компоненты $E_к$ и $R_к$ не оказывают влияния на ток коллектора, поскольку они включены последовательно с генераторами тока. Величины $E_к$ и $R_к$ связаны соотношениями (рис. 6-2, а):

$$-E_к = -I_к R_к + U_{кб} + U_б \quad (6-5а)$$

или

$$-E_к = -I_к R_к + U_{кэ} + U_э. \quad (6-5б)$$

* Для кремниевых транзисторов обычно принимают $U_{эб} = 0,7$ В, см. с. 146.

Здесь ток I_k и напряжения $U_{кб}$ и $U_{кэ}$ являются заданными координатами рабочей точки, а потенциалы $U_б$ и $U_э$ можно записать в виде

$$U_б = -E_б + I_б R_б; \quad U_э = E_э - I_э R_э,$$

где токи $I_б$, $I_э$ — однозначные функции тока I_k [см. (4-8) и (4-72)].

Выражения (6-4в) и (6-5) позволяют выбрать значения $E_э$, $E_б$ и E_k , а также $R_э$, $R_б$ и R_k , если известны параметры транзистора (β , $I_{к0}$) и его рабочая точка ($I_{кА}$, $U_{кА}$). При этом, разумеется, некоторыми из шести искомых величин заранее задаются, так как они связаны всего двумя уравнениями¹.

Как правило, коллекторная цепь рассчитывается, исходя из соотношений (6-5), а цепи смещения — исходя из выражений (6-4). Дополнительные критерии для выбора значений $E_{эб}$, $R_э$ и $R_б$ будут выведены в следующем параграфе.

6-2. СТАБИЛЬНОСТЬ РАБОЧЕЙ ТОЧКИ

Всякое смещение рабочей точки, характеризуемое приращениями $\Delta I_{кА}$ и $\Delta U_{кА} = \Delta I_{кА} R_k$, вызывает изменение дифференциальных параметров транзистора, поскольку они зависят от режима (см. § 4-6). Если смещение точки A обусловлено изменением температуры (см. точку A' на рис. 6-1), получается так называемая косвенная зависимость параметров от температуры². Большие смещения $\Delta I_{кА}$ и $\Delta U_{кА}$ могут привести к существенным нелинейным искажениям и даже к частичной или полной отсечке сигнала (например, если точка A' попадает в область насыщения). Поэтому стабилизация рабочей точки с помощью того или иного вида обратной связи по постоянному току является одной из важнейших задач при проектировании транзисторных каскадов [112, 113].

На рабочий ток транзистора, а значит, и на стабильность рабочей точки влияют следующие основные причины: тепловой ток $I_{к0}$, напряжение на эмиттерном переходе $U_{эб}$ и интегральный коэффициент передачи тока β . Все они зависят от температуры, а также подвержены временному дрейфу. Влиянием сопротивления r_k мы пренебрежем по тем же соображениям, что и при анализе постоянных составляющих (см. сноску на с. 318).

Общий анализ. Полное приращение коллекторного тока имеет вид:

$$\Delta I_k = \frac{\partial I_k}{\partial I_{к0}} \Delta I_{к0} + \frac{\partial I_k}{\partial U_{эб}} \Delta U_{эб} + \frac{\partial I_k}{\partial \beta} \Delta \beta.$$

Это приращение можно конкретизировать, дифференцируя выражение (6-4в) по соответствующим параметрам. Однако проще

¹ В конкретных схемах некоторые э. д. с. и резисторы могут отсутствовать (см., например, рис. 6-1).

² Термин «косвенная» призван отличать эту зависимость от «прямой», которая имеет место при неизменной рабочей точке (см. рис. 4-20).

поступить следующим образом. Исходя из формулы (4-72), запишем:

$$\Delta I_k = \Delta \beta I_6 + \beta \Delta I_6 + (1 + \beta) \Delta I_{k0} + \Delta \beta I_{k0}. \quad (6-6a)$$

Подставим сюда приращение ΔI_6 , выраженное с помощью (6-4a):

$$\Delta I_6 = -\frac{\Delta U_{эб}}{R_{эб}} - \gamma_6 \Delta I_k \quad (6-6б)$$

и решим уравнение относительно ΔI_k . Тогда

$$\Delta I_k = \frac{\beta}{1 + \beta \gamma_6} \left[\frac{\Delta I_{k0}}{\alpha} - \frac{\Delta U_{эб}}{R_{эб}} + (I_6 + I_{k0}) \frac{\Delta \beta}{\beta} \right].$$

Величина, вынесенная за квадратные скобки, носит название *коэффициента неустойчивости*¹ и обычно обозначается буквой S :

$$S = \frac{\beta}{1 + \beta \gamma_6}. \quad (6-7)$$

Используя это обозначение, а также учитывая, что $I_6 + I_{k0} = I_6 / (1 + \beta)$, запишем полное приращение ΔI_k следующим образом:

$$\Delta I_k = S \left(\frac{\Delta I_{k0}}{\alpha} - \frac{\Delta U_{эб}}{R_{эб}} + \frac{I_6}{1 + \beta} \frac{\Delta \beta}{\beta} \right). \quad (6-8)$$

Выражениям (6-6) соответствует эквивалентная схема каскада, показанная на рис. 6-2, б и действительная только для приращений токов и напряжений относительно их начальных значений. На рис. 6-2, б каждая причина неустойчивости отражена соответствующим генератором: $\Delta U_{эб}$, $(1 + \beta) \Delta I_{k0}$ и $(I_{k0} + I_6) \Delta \beta$, где I_6 — ток базы в рабочей точке.

Очевидно, что приращение ΔI_k будет тем меньше, чем меньше коэффициент неустойчивости. Из формулы (6-7) получаются предельные значения:

$$S_{\min} = \alpha; \quad S_{\max} = \beta,$$

первое из которых соответствует значению $\gamma_6 = 1$, а второе — значению $\gamma_6 = 0$.

Следовательно, для получения максимальной стабильности нужно стремиться к выполнению условия $\gamma_6 \approx 1$ или вытекающего из него неравенства

$$R_э \gg R_6. \quad (6-9)$$

Условие (6-9) служит надежным ориентиром при проектировании стабильных транзисторных каскадов, однако выполнение его далеко не всегда возможно и необходимо. Часто вполне удов-

¹ Первоначально этот коэффициент был введен в практику в виде производной dI_k/dI_{k0} и записывался в более частной форме [113]. Такое определение неполно, а в ряде случаев (особенно для кремниевых транзисторов, у которых током I_{k0} обычно пренебрегают) оно совершенно неприемлемо.

летворительные результаты дают значения $R_9/R_6 = 0,5 \div 1$, которыми соответствуют значения $\gamma_6 = 0,3 \div 0,5$ и $S = 2 \div 3$.

В общем виде связь между R_9/R_6 и S легко получается из выражений (6-7) и (6-4б):

$$\frac{R_9}{R_6} = \frac{\beta - S}{\beta(S-1) + S} \approx \frac{1}{S-1}. \quad (6-10)$$

Принятое здесь приближение оправдывается при $\beta \gg S$.

Если поделить (6-8) на (6-4в), то с учетом (6-7) относительное изменение коллекторного тока будет иметь вид:

$$\frac{\Delta I_K}{I_K} = \frac{-\Delta U_{эб} + \left[\frac{\Delta I_{к0}}{\alpha} + \frac{I_9}{1+\beta} \frac{\Delta \beta}{\beta} \right] R_{эб}}{E_{эб} + I_{к0} R_{эб} - U_{эб}}. \quad (6-11)$$

Как видим, относительное изменение тока не зависит от коэффициента неустойчивости (т. е. от отношения R_9/R_6), но находится в прямой зависимости от суммарного сопротивления $R_{эб}$; поэтому последнее нужно делать достаточно малым. В пределе, при $R_{эб} \rightarrow 0$, относительное изменение тока определяется относительной неустойчивостью напряжения $U_{эб}$.

Выражение (6-8) или (6-11) вместе с выражением (6-4в) обеспечивает две связи между тремя величинами $E_{эб}$, R_9 и R_6 , подлежащими определению при статическом расчете каскада. Поэтому, задаваясь одной из этих величин, можно найти две другие.

До сих пор мы использовали приращения $\Delta I_{к0}$, $\Delta U_{эб}$ и $\Delta \beta$, не оговаривая, какими причинами они обусловлены: разбросом параметров от транзистора к транзистору, временной «ползучестью» или температурным дрейфом. На практике чаще всего приходится учитывать влияние температуры. Для этого, зная температурный диапазон, рассчитывают значения $\Delta I_{к0}$ и $\Delta U_{эб}$ соответственно по формулам (2-46) и (2-66), (2-67), а значение $\Delta \beta$ находят либо из графиков, приводимых в справочниках, либо специально измеряют.

В случае кремниевых транзисторов формулы (6-4), (6-8) и (6-11) существенно упрощаются, потому что для качественных кремниевых транзисторов вплоть до температур $+100 \div +125^\circ \text{C}$ можно пренебречь величинами $I_{к0}$ и $\Delta I_{к0}$. Это, однако, не значит, что кремниевые транзисторы обеспечивают более высокую температурную стабильность каскадов, так как им часто присущи более сильные зависимости $\beta(T)$ и $U_{эб}(T)$, чем германиевым транзисторам.

Оценим относительную роль тех причин температурной неустойчивости, которые характеризуются отдельными слагаемыми в формуле (6-8). Пусть рабочий диапазон температур составляет $\pm 50^\circ \text{C}$. В этом диапазоне изменение теплового тока для германиевых транзисторов может составить $\Delta I_{к0} \approx 40 \text{ мкА}$ (для кремниевых транзисторов значение $\Delta I_{к0}$ будет примерно на два порядка меньше, и им можно пренебречь). Второе слагаемое найдем, полагая $R_{эб} \approx 3 \text{ кОм}$ и $\Delta U_{эб} \approx \varepsilon \Delta T$, где $\varepsilon = -1,5 \text{ мВ/}^\circ \text{C}$; тогда $\varepsilon \Delta T / R_{эб} = 50 \text{ мкА}$. Для более точного вычисления третьего слагаемого целесообразно считать, что один из коэффициентов β , стоящих в знаменателе, соответствует нижней, а другой — верхней границе температурного диапазона. Пусть, например, $I_9 = 2 \text{ мА}$; $\beta_{\text{мин}} = 30$; $\beta_{\text{макс}} = 90$;

тогда $\Delta\beta = 60$ и $I_{\text{э}}\Delta\beta/(1 + \beta_{\text{макс}})\beta_{\text{мин}} \approx 44$ мкА. Как видим, все приращения в нашем примере равноценны. Следует, однако, иметь в виду, что роль напряжения $U_{\text{эб}}$ возрастает с уменьшением сопротивлений, а роль коэффициента β — с увеличением токов.

Стабильность типовых схем. Как уже отмечалось, наибольшее значение в усилителях имеют каскады ОЭ. Поэтому им будет уделено главное внимание и в данном разделе.

Однако начнем с более простого каскада ОБ (рис. 6-3). Структура его вполне соответствует обобщенной схеме (см. рис. 6-2), если под сопротивлением $R_{\text{Б}}$ понимать только объемное сопротивление базы $r_{\text{б}}$. При токах эмиттера до 10—50 мА, легко выполняется условие (6-9) и стабильность будет весьма высокой. При токах в сотни

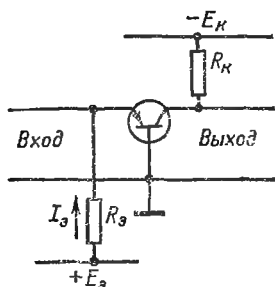


Рис. 6-3. Схема, обеспечивающая режим каскада ОБ.

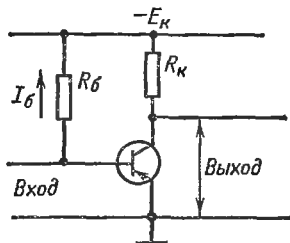


Рис. 6-4. Простейшая схема, обеспечивающая режим каскада ОЭ.

миллиампер и выше $R_{\text{э}}$ должно быть малым, а так как сопротивление $r_{\text{б}}$ — величина заданная, то отношение $R_{\text{э}}/r_{\text{б}}$ уменьшается и стабильность каскада падает.

Простейшая схема каскада ОЭ показана на рис. 6-4. В ней базовый ток задается резистором $R_{\text{б}}$, а сопротивление $R_{\text{э}} = 0$. Поэтому $\gamma_{\text{б}} = 0$ и стабильность данного каскада оказывается очень низкой ($S = \beta$).

На рис. 6-5 показана типичная схема каскада ОЭ, наиболее часто используемая на практике. Сравнивая ее с обобщенной схемой (см. рис. 6-2), приходим к выводу, что в данном случае $R_{\text{б}} = R_1 \parallel R_2$. Принцип стабилизации в схеме на рис. 6-5 заключается в том, что делитель R_1, R_2 задает потенциал базы и тем самым достаточно жестко фиксирует потенциал эмиттера, поскольку $U_{\text{эб}} \approx \text{const}$. В этих условиях ток $I_{\text{э}}$, равный $U_{\text{э}}/R_{\text{э}}$, не может сильно меняться. Чем меньше сопротивления делителя, тем меньше зависит потенциал базы от изменений базового тока и тем лучше стабилизация. Однако очень низкоомный делитель R_1, R_2 невыгоден по двум причинам: возрастает расход мощности от источника питания и шунтируется входное сопротивление каскада. Поэтому обычно делают $R_1 \parallel R_2 \approx \approx R_{\text{э}}$ или больше. Получаемые при этом значения $S = 2 \div 5$ бес-

печивают абсолютную стабильность, в 5—10 раз более высокую, чем в схеме на рис. 6-4. Сопротивление R_3 желательно выбирать большим, так как при этом можно повысить сопротивление делителя. Однако величина R_3 ограничена ростом падения напряжения $I_3 R_3$.

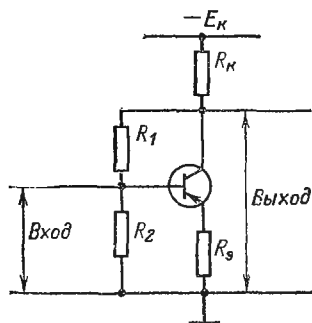


Рис. 6-5. Типичная схема, обеспечивающая стабильный режим каскада ОЭ.

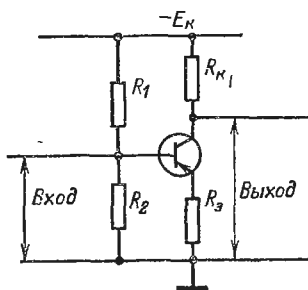


Рис. 6-6. Вариант схемы, обеспечивающий повышенную стабильность режима в каскаде ОЭ.

На рис. 6-6 показана схема, в которой в отличие от схемы на рис. 6-5 напряжение, питающее делитель, меняется при изменении потенциала U_k , а значит, и тока I_k . Увеличение последнего уменьшает (по модулю) потенциал U_k и соответственно потенциал базы. Поэтому результирующее изменение ΔI_k получается меньше,

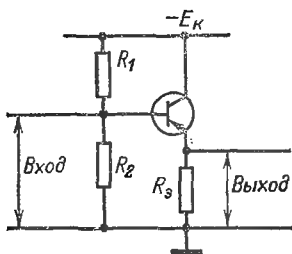


Рис. 6-7. Схема, обеспечивающая режим каскада ОК.

чем в схеме на рис. 6-5. Анализ показывает, что этот выигрыш тем более заметен, чем меньше отношение $(R_1 + R_2)/R_k$. Преимущества данной схемы проявляются тогда, когда указанное отношение близко к единице или меньше ее. Однако такой низкоомный делитель, улучшая стабильность в 2—3 раза¹, заметно снижает усиление каскада и входное сопротивление.

Каскад ОК обычно имеет структуру, показанную на рис. 6-7. Она отличается от структуры каскада ОЭ (см. рис. 6-5) отсутствием коллекторного резистора R_k . Так как сопротивление последнего не влияет на приращения ΔI_k , то формула (6-8) остается в силе. Для каскада ОК характерно то, что приращения ΔI_k меньше влияют на величину коллекторного напряжения из-за отсутствия R_k . Следовательно, при малых R_3 будет меняться в основном одна координата рабочей точки (I_k) и параметры транзистора будут меняться лишь постольку, поскольку они зависят от I_k .

¹ При этом возможны значения $S < \alpha$ (см. гл. 4 в [85]).

Необходимо заметить, что возможна большая стабильность рабочей точки не всегда полезна. Действительно, параметры транзистора зависят от температуры не только косвенно, из-за смещения рабочей точки, но и непосредственно, т. е. если даже рабочая точка неизменна. Например, коэффициент β увеличивается с ростом температуры при постоянных токе I_3 и напряжении U_K . Поэтому иногда допускают некоторое увеличение тока и уменьшение напряжения с тем, чтобы режимное приращение $\Delta\beta(I_3; U_K)$ частично компенсировало непосредственное температурное приращение $\Delta\beta(T)$. Тем самым будет достигнута большая стабильность усиления, если именно она является первоочередным фактором.

6-3. РАСЧЕТ КАСКАДОВ ПО ПОСТОЯННОМУ ТОКУ

Будем считать заданными режимные величины $I_3 \approx I_K$ и $U_{K6} \approx U_{K3}$. Они либо заимствуются из справочника, либо выбираются из соображений, кратко сформулированных в § 6-1. Кроме того, можно считать известными сопротивления $R_{вх}$ и R_K , которые получаются в результате предшествующего расчета каскада по переменному току. Из параметров транзистора следует знать β (или α) и его зависимость от температуры, тепловой ток I_{K0} и сопротивление базы r_6 . Наконец, должны быть заданы рабочий диапазон температуры и допустимый сдвиг рабочей точки в этом диапазоне ($\Delta I_{K, \text{доп}}$ или $\Delta I_{K, \text{доп}}/I_K$). Тогда для статического расчета каскадов можно руководствоваться следующими соображениями.

Каскад с общей базой. В этом каскаде (см. рис. 6-3) имеем $E_{36} = E_3$; $R_6 = 0$ и соответственно $R_{36} = R_3$ и $\gamma_6 = 1$.

Следовательно, выражения (6-4в) и (6-8) [или (6-11)] содержат только две неизвестные величины: R_3 и E_3 , которые поэтому определяются однозначно.

Полагая $\Delta I_K \leq \Delta I_{K, \text{доп}}$, из (6-8) получаем:

$$R_3 \geq \frac{-\alpha \Delta U_{36}}{\Delta I_{K, \text{доп}} - \Delta I_{K0} - \frac{I_3}{1 + \beta} \frac{\Delta \beta}{\beta}}. \quad (6-12)$$

После того как R_3 выбрано из этого условия, выражение (6-4в) с учетом (4-8) приводит к соотношению

$$E_3 = I_3 R_3 + U_{36}, \quad (6-13)$$

которое можно получить и непосредственно из принципиальной схемы каскада.

Нередко найденное значение E_3 оказывается столь малым, что его трудно получить в реальной схеме (обычно желательно иметь $E_3 > 1 \div 2$ В). Поэтому, если из формулы (6-13) получается слишком малое значение E_3 , задаются более приемлемым значением Э. д. с. и соответственно увеличивают R_3 . Это способствует выполнению неравенства

$$R_3 \geq R_{вх}, \quad (6-14)$$

которое желательно с той точки зрения, чтобы сопротивление R_3 возможно меньше шунтировало вход каскада.

Коллекторная цепь рассчитывается из соотношения

$$E_k = |U_{кб}| + I_k R_k, \quad (6-15)$$

где напряжение $U_{кб}$ заранее известно, а одним из параметров E_k или R_k задаются. Соотношение (6-15) получается из общей формулы (6-5а) при $U_6 = 0$.

Каскад с общим эмиттером. Рассмотрим сначала типичную схему (рис. 6-5), в которой $E_{эб} = E_6 = E_k R_2 / (R_1 + R_2)$ и $R_6 = R_1 \parallel R_2$. Поскольку в уравнения (6-4в) и (6-8) [или (6-11)] входят три неизвестные величины: E_6 , R_6 и R_9 , одной из них нужно задаться. Часто задаются сопротивлением R_6 из условия, аналогичного (6-14):

$$R_6 = R_1 \parallel R_2 \gg R_{вх}, \quad (6-16)$$

но иногда разумнее принять за основу сопротивление R_9 , исходя из желательного значения напряжения $U_9 = I_9 R_9$.

В первом случае, выбрав R_6 , находим R_9 с помощью выражения (6-8), полагая $\Delta I_k \leq \Delta I_{к. доп}$:

$$R_9 \geq \frac{\alpha (\Delta I_{к. доп} R_6 - \Delta U_{эб})}{\Delta I_{к. доп} - \left[\Delta I_{к0} + \frac{I_9 \Delta \beta}{(1 + \beta)^2} \right]} - R_6. \quad (6-17a)$$

Во втором случае, выбрав R_9 , находим R_6 :

$$R_6 \leq \frac{\beta (\Delta I_{к. доп} R_9 + \Delta U_{эб})}{(1 + \beta) \left[\Delta I_{к0} + \frac{I_9 \Delta \beta}{(1 + \beta)^2} \right] - \Delta I_{к. доп}} - R_9; \quad (6-17б)$$

если знаменатель (6-17б) отрицательный, следует изменить знак неравенства.

Задаваясь значением $\Delta I_{к. доп}$, следует обеспечивать условие

$$\Delta I_{к. доп} \geq \Delta I_{к0} + I_9 \frac{\Delta \beta}{(1 + \beta)^2}; \quad (6-18)$$

в противном случае формулы (6-17) недействительны. Иначе говоря, значение $\Delta I_{к. доп}$, не удовлетворяющую условию (6-18), нельзя реализовать ни при каких значениях R_9 и R_6 .

Зная R_9 и R_6 , нетрудно найти сопротивления делителя R_1 и R_2 . Для этого запишем потенциал базы в виде

$$U_6 = -E_6 + I_6 R_6 = -E_k \frac{R_6}{R_1} + I_6 R_6$$

и подставим $U_6 = -I_9 R_9 - U_{эб}$. Тогда получим:

$$R_1 = \frac{E_k}{\frac{I_9 R_9 + U_{эб}}{R_6} + I_6}. \quad (6-19a)$$

Сопротивление R_2 легко выражается через R_6 и R_1 :

$$R_2 = \left(\frac{1}{R_6} - \frac{1}{R_1} \right)^{-1}. \quad (6-196)$$

Из общей формулы (6-56) при $E_9 = 0$ получаем соотношение

$$E_k = |U_{k9}| + I_k R_k + I_9 R_9, \quad (6-20)$$

исходя из которого можно определить э. д. с. E_k или (если э. д. с. E_k задана) внести необходимые коррективы в значения R_k и R_9 (последнее с точки зрения стабильности желательно изменять только в сторону увеличения).

Если каскад имеет структуру, показанную на рис. 6-6, то сопротивлением R_k и напряжением E_k задаваться не следует, так как в противном случае эта схема не будет иметь преимуществ перед предыдущей. Сопротивления R_6 и R_9 выбираются с учетом условий (6-16) и (6-17). После этого нужно найти сопротивления R_1 и R_2 .

Сопротивление R_1 определяется из выражения (6-19а), если вместо э. д. с. E_k использовать напряжение $U_k = I_9 R_9 + |U_{k9}|$ (от которого питается делитель в данной схеме):

$$R_1 = \frac{|U_{k9}| + I_9 R_9}{\frac{I_9 R_9 + U_{96}}{R_6} + I_6}. \quad (6-21)$$

Сопротивление R_2 рассчитывается по формуле (6-196). Зная сумму $R_1 + R_2$, выбирают R_k из условия

$$R_k \geq R_1 + R_2. \quad (6-22)$$

Учитывая, что через сопротивление R_k протекают и ток коллектора, и ток делителя, напряжение E_k определяют по формуле

$$E_k = |U_{k9}| + I_9 R_9 + \left(I_k + \frac{|U_{k9}| + I_9 R_9}{R_1 + R_2} \right) R_k. \quad (6-23)$$

В заключение заметим, что схемы на рис. 6-5 и 6-6 в действительности должны быть дополнены блокирующими конденсаторами. В противном случае усиление, даваемое каскадами, будет очень мало, так как обратные связи будут подавлять не только температурные приращения ΔI_k , но и приращения, вызванные сигналом. Реальные схемы каскадов будут рассмотрены в следующей главе.

Каскад с общим коллектором. Как уже отмечалось, этот каскад (рис. 6-7) с точки зрения режима отличается от каскада ОЭ только отсутствием резистора R_k . Поэтому расчет сопротивлений проводится по формулам (6-17) и (6-19). Учитывая, что $R_k = 0$, значение R_9 согласно (6-20) можно выбрать значительно большим, чем в каскаде ОЭ. Это способствует повышению стабильности, а главное, позволяет увеличить сопротивление R_6 и вместе с ним входное сопротивление каскада.

УСИЛИТЕЛИ С ЕМКОСТНОЙ СВЯЗЬЮ

7-1. ВВЕДЕНИЕ

В транзисторных усилителях, как и в ламповых, наибольшее распространение имеют каскады с емкостной связью¹.

Типичный каскад такого типа (рис. 7-1) содержит переходные конденсаторы C_1 и C_2 и блокирующий конденсатор C_3 .

Конденсатор C_1 призван изолировать генератор (источник) сигнала от входа каскада по постоянному току и соединить их по переменной составляющей сигнала. Конденсатор C_2 выполняет аналогичную функцию по отношению к выходу каскада и нагрузке. Конденсатор C_3 шунтирует резистор R_3 по переменному току и тем самым ослабляет (практически ликвидирует) отрицательную обратную связь в каскаде на частоте сигнала.

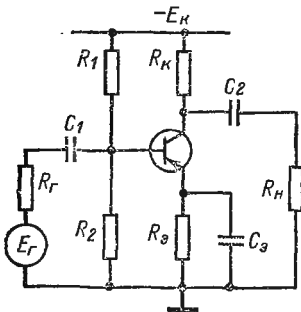


Рис. 7-1. Принципиальная схема каскада ОЭ с емкостной связью.

Емкости этих конденсаторов оказывают влияние на работу каскада в области низших частот и при передаче вершины импульсов. В области высших частот и при передаче фронтов импульсов первостепенное значение имеют временная (или частотная) зависимость коэффициента β и коллекторная емкость C_k . В отличие от ламповых схем паразитные емкости в транзисторных каскадах оказывают обычно второстепенное влияние и могут не учитываться.

Как и в других типах усилителей, при анализе транзисторных каскадов выделяют область средних частот, в которой можно считать схему состоящей только из активных элементов и не учитывать временных и частотных искажений. С этой области мы и начнем изучение каскада.

Однако предварительно сделаем два замечания.

Входные сопротивления биполярных транзисторов, как правило, меньше, чем выходные. В связи с этим транзисторные каскады часто работают не как усилители напряжения, а как усилители тока [62]. Поэтому наряду с коэффициентом усиления

¹ Ниже термин «каскад» (не совсем строго) используется применительно к однокаскадным усилителям — главным объектам анализа в данной части книги. Собственно каскады, т. е. элементарные промежуточные ячейки многокаскадного усилителя (см. рис. 7-16), являются частным случаем однокаскадных усилителей, когда источником сигнала служит выход предыдущего каскада, а нагрузкой — вход следующего каскада (см. § 7-4).

Что касается тезиса о распространенности емкостной связи, то он справедлив лишь для дискретных электронных схем, монтируемых из отдельных компонентов. В области интегральных микроэлектронных схем типичны гальванические, непосредственные связи между каскадами (см. гл. 13 и 14).

напряжения K_u полезно дополнительно анализировать коэффициент усиления тока K_i , хотя они и связаны друг с другом (см. гл. 8).

Для облегчения анализа целесообразно упростить эквивалентную схему транзистора, пренебрегая генератором обратной связи $\mu_{\text{ЭК}} U_K$. Такое упрощение предполагает соблюдение неравенства

$$\mu_{\text{ЭК}} U_K \ll U_{\text{бэ}}.$$

Заменяя U_K на $U_{\text{вых}}$ и $U_{\text{бэ}}$ на $U_{\text{вх}}$, представим это неравенство в более наглядном виде:

$$K_u \ll \frac{1}{\mu_{\text{ЭК}}}. \quad (7-1)$$

Практически у всех транзисторов $\mu_{\text{ЭК}} < 10^{-3}$; следовательно, обратная связь по напряжению пренебрежимо мала при $K_u < 100$, т. е. в большинстве реальных каскадов.

7-2. КАСКАД В ОБЛАСТИ СРЕДНИХ ЧАСТОТ

В области средних частот внешние емкости C_1 , C_2 и C_3 будем предполагать бесконечно большими, емкость C_K равной нулю, сопротивления нагрузки R_H и источника сигнала R_r

чисто активными, а коэффициент передачи β действительной величиной. Тогда эквивалентная схема каскада будет такой, как показано на рис. 7-2. В этой схеме отсутствует сопротивление r_K^* (оно будет учтено при более полном анализе), а эквивалентное сопротивление делителя $R_1 \parallel R_2$ показано пунктиром (его роль также будет рассмотрена несколько позднее). Займемся сначала анализом основной части схемы, показанной сплошными линиями.

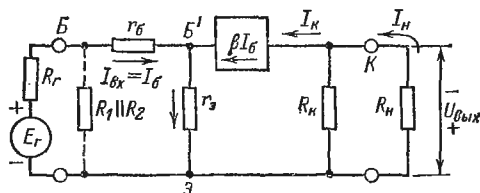


Рис. 7-2. Упрощенная эквивалентная схема каскада в области средних частот.

Упрощенный анализ. Входное сопротивление каскада определяется отношением $U_{\text{вх}}/I_{\text{вх}}$, где $U_{\text{вх}}$ — напряжение на зажимах Б — Э, а $I_{\text{вх}}$ — ток базы. Учитывая, что через сопротивление $r_б$ протекает ток $I_б$, а через сопротивление $r_э$ — ток $(1 + \beta)I_б$, и складывая оба падения напряжения, находим величину $U_{\text{вх}}$; деля ее на ток $I_б = I_{\text{вх}}$, получаем:

$$R_{\text{вх}} = r_б + (1 + \beta) r_э. \quad (7-2)$$

Если, например, $r_б = 100$ Ом, $r_э = 25$ Ом и $\beta = 50$, то $R_{\text{вх}} = 1350$ Ом.

Входное сопротивление зависит от режима транзистора, особенно от тока $I_э$, поскольку величина $r_э$ обратно пропорциональна этому току. При токах $I_э > 10$ мА основное зна-

чение имеет сопротивление базы; оно определяется минимальное значение $R_{\text{вх}}$. Повышение входного сопротивления, желательное в ряде случаев, достигается уменьшением эмиттерного тока в рабочей точке. Однако при этом уменьшается усиление каскада и могут возрасти нелинейные искажения, так как в области малых токов входная характеристика имеет большую кривизну (см. рис. 4-11, б).

Выходное сопротивление, как всегда, определяется со стороны выходных зажимов при отключенной нагрузке и нулевом входном сигнале. Если не учитывать величины $r_{\text{к}}$ (рис. 7-2), то выходное сопротивление равно сопротивлению $R_{\text{к}}$:

$$R_{\text{вых}} = R_{\text{к}}. \quad (7-3)$$

О выборе величины $R_{\text{к}}$ будет сказано ниже.

Коэффициент усиления напряжения определим как отношение напряжения $U_{\text{вых}}$ на нагрузке к э. д. с. источника сигнала $E_{\text{г}}$ *. Из очевидных соотношений

$$U_{\text{вых}} = -\beta I_{\text{б}} (R_{\text{к}} \parallel R_{\text{н}});$$

$$I_{\text{б}} = \frac{E_{\text{г}}}{R_{\text{г}} + R_{\text{вх}}}$$

легко находим:

$$K_U = \frac{U_{\text{вых}}}{E_{\text{г}}} = -\frac{\beta (R_{\text{к}} \parallel R_{\text{н}})}{R_{\text{г}} + R_{\text{вх}}}. \quad (7-4a)$$

В идеальном усилителе напряжения ($R_{\text{г}} = 0$), работающем в режиме холостого хода ($R_{\text{н}} = \infty$), коэффициент усиления будет максимальным и равным:

$$K_{u \text{ х. х}} = -\frac{\beta R_{\text{к}}}{R_{\text{вх}}} \approx -\frac{R_{\text{к}}}{r_{\text{э}}}. \quad (7-4б)$$

Введя так называемые коэффициенты входа и выхода каскада по напряжению

$$\xi_{\text{вх}u} = \frac{R_{\text{вх}}}{R_{\text{г}} + R_{\text{вх}}}; \quad \xi_{\text{вых}u} = \frac{R_{\text{н}}}{R_{\text{вых}} + R_{\text{н}}} = \frac{R_{\text{н}}}{R_{\text{к}} + R_{\text{н}}},$$

можно записать (7-4a) в следующей форме:

$$K_U = K_{u \text{ х. х}} \xi_{\text{вх}u} \xi_{\text{вых}u}. \quad (7-4в)$$

В дальнейшем будем использовать только индекс u , считая, однако, что величина K_u определяется по отношению к $E_{\text{г}}$. В тех немногих случаях, когда величина K_u относится к $U_{\text{вх}}$, будем делать соответствующие оговорки. Если в формулах фигурируют коэффи-

* При сравнимых значениях $R_{\text{г}}$ и $R_{\text{вх}}$ такое определение более логично, чем то, которое применялось в ламповых каскадах (отношение $U_{\text{вых}}$ к напряжению непосредственно на сетке лампы $U_{\text{вх}}$). Действительно, для ламп, как правило, имеет место соотношение $R_{\text{г}} \ll R_{\text{вх}}$ и, следовательно, $U_{\text{вх}} \approx E_{\text{г}}$, тогда как в случае транзисторов, как правило, $R_{\text{г}} \lesssim R_{\text{вх}}$ и потому $U_{\text{вх}} < E_{\text{г}}$. Поскольку первичным сигналом является э. д. с. $E_{\text{г}}$, именно по отношению к ней и следует в общем случае определять коэффициент усиления.

пциенты усиления обоих типов каскадов, например (7-4в), то для величин, отнесенных к э. д. с. E_r , используется индекс U , а для величин, отнесенных к напряжению $U_{вх}$, — индекс u .

В выражениях (7-4) нужно обратить внимание на знак минус, который говорит об изменении полярности сигнала (т. е. сдвиге фазы на 180°). Иногда для простоты знак минус не пишут, но о нем следует всегда помнить.

Из формул (7-4) видно, что коэффициент усиления увеличивается с уменьшением входного сопротивления. Это обстоятельство специфично для транзисторных каскадов и объясняется тем, что выходное напряжение пропорционально входному току, а он растет с уменьшением $R_{вх}$.

Далее формула (7-4а) показывает, что суммарную коллекторную нагрузку $R_k \parallel R_n$ желательно делать возможно большей. Так как сопротивление R_n обычно задано, следует выбирать $R_k > R_n$. В частности, если нагрузкой является вход аналогичного каскада, то $R_n = R_{вх}$ и R_k определяется из неравенства

$$R_k > R_{вх}, \quad (7-5)$$

которое должно быть достаточно сильным.

Вообще говоря, при малой величине R_r и большой величине $R_k \parallel R_n$ коэффициент усиления может быть весьма высоким и ограничен факторами, которые мы отметим ниже, учтя сопротивление r_k^* . Однако в типичном случае, когда данный каскад работает на другой аналогичный каскад, значение K_u ограничено. Пусть, например, $R_k \gg R_n = R_{вх}$ и $R_r \ll R_{вх}$. Тогда из (7-4а) получаем $K_u \approx \beta$. Пусть теперь $R_k = R_r = R_{вх}$; тогда $K_u = \beta/4$. Если β имеет значение 30—100, коэффициент усиления будет соответственно лежать в пределах от $K_u \approx 30 \div 100$ до $K_u \approx 8 \div 25$.

Теперь учтем влияние базового делителя, сопротивление которого не всегда может считаться пренебрежимо большим. Воспользуемся теоремой об эквивалентном генераторе [62] и с ее помощью преобразуем источник сигнала E_r , R_r в эквивалентный источник с параметрами:

$$E_r' = E_r \frac{R_1 \parallel R_2}{R_r + R_1 \parallel R_2}; \quad R_r' = R_r \parallel R_1 \parallel R_2. \quad (7-6)$$

Легко убедиться, что выходное напряжение при этом уменьшится, несмотря на некоторое уменьшение E_r' по сравнению с E_r . Поскольку величину $U_{вых}$ нужно относить к прежнему, «истинному» значению E_r , приходим к выводу, что наличие делителя (если его сопротивление $R_1 \parallel R_2$ сравнимо со значением R_r) снижает усиление каскада.

Коэффициент усиления тока можно найти, воспользовавшись принципом взаимного соответствия [62], т. е. заменив генератор э. д. с. E_r , включенный последовательно с сопротивлением R_r , генератором тока I_r , включенным параллельно с сопротивлением R_r . Тогда ток базы и ток в нагрузке запишутся следующим образом:

$$I_6 = I_r \frac{R_r}{R_r + R_{вх}}; \quad I_n = -\beta I_6 \frac{R_k}{R_k + R_n}.$$

Отсюда легко получаем:

$$K_I = \frac{I_n}{I_r} = -\beta \frac{R_r}{R_r + R_{вх}} \frac{R_k}{R_k + R_n}. \quad (7-7a)$$

В идеальном усилителе тока ($R_r = \infty$), работающем в режиме короткого замыкания ($R_n = 0$), имеем:

$$K_{i_{к.з}} = -\beta. \quad (7-7б)$$

Выражение (7-7a) можно записать в виде

$$K_I = K_{i_{к.з}} \xi_{вх} \xi_{вых}, \quad (7-7в)$$

где

$$\xi_{вх} = \frac{R_r}{R_r + R_{вх}}; \quad \xi_{вых} = \frac{R_{вых}}{R_{вых} + R_n} = \frac{R_k}{R_k + R_n}$$

— коэффициенты входа и выхода каскада по току. Относительно индексов I и i будем придерживаться правил, аналогичных тем, которые изложены выше в связи с формулами (7-4).

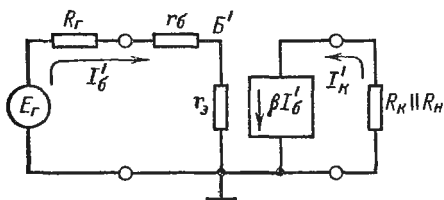


Рис. 7-3. Случай отсутствующей внутренней обратной связи по току.

Внутренняя обратная связь. Из рис. 7-2 ясно, что в рассматриваемом каскаде имеет место обратная связь, поскольку часть коллекторного тока ответвляется во входную цепь. Для того чтобы исключить обратную связь, можно мысленно отсоединить генератор βI_6 от точки B' и заземлить его левый зажим (рис. 7-3); тогда выходной ток I'_k не будет влиять на ток источника сигнала (более строго вопросы внутренней обратной связи рассматриваются в § 8-3).

Коэффициент обратной связи в реальной схеме (рис. 7-2) определяется соотношением

$$\gamma_6 = \frac{I_6}{I_k}, \quad (7-8)$$

где I_k — ток, задаваемый в узловую точку B' от независимого генератора тока, заменяющего генератор βI_6 .

Элементарный расчет показывает, что в схеме на рис. 7-2 коэффициент обратной связи равен:

$$\gamma_6 = \frac{r_3}{R_r + r_6 + r_3}. \quad (7-9)$$

Аналогичная величина, действительная для полных токов и зависящая от внешних сопротивлений, встречалась при анализе температурной стабильности [см. (6-36), (6-46)].

Ниже нередко придется иметь дело с частным случаем выражения (7-9), когда $R_r = 0$ и токораспределение обусловлено только

сопротивлениями самого транзистора:

$$\gamma_{60} = \frac{r_9}{r_6 + r_9}. \quad (7-10)$$

Заметим еще, что доля коллекторного тока, ответвляющаяся в цепь эмиттера, связана с коэффициентом γ_6 элементарным соотношением

$$\gamma_9 = 1 - \gamma_6. \quad (7-11)$$

В предыдущем разделе параметры каскада ОЭ были получены непосредственно, но их же можно получить, учитывая наличие обратной связи, и представить в соответствующей форме. Для этого воспользуемся схемой на рис. 7-3 и запишем сначала параметры без обратной связи¹:

$$R_{вх} = r_6 + r_9; \quad (7-12)$$

$$K_u = - \frac{\beta (R_k \parallel R_n)}{R_r + r_6 + r_9}; \quad (7-13)$$

$$K_i = - \beta \frac{R_r}{R_r + r_6 + r_9} \frac{R_k}{R_k + R_n}. \quad (7-14)$$

Сравнивая выражения (7-12)—(7-14) с выражениями (7-2), (7-4а) и (7-7а), нетрудно убедиться, что они связаны вполне определенными соотношениями. А именно, входное сопротивление (7-2) получается, если умножить (7-12) на $1 + \beta\gamma_{60}$:

$$R_{вх} = (r_6 + r_9) (1 + \beta\gamma_{60}), \quad (7-15)$$

а коэффициенты усиления (7-4а) и (7-7а) получаются путем деления выражений (7-13) и (7-14) на $1 + \beta\gamma_6$:

$$K_u = - \frac{\beta}{1 + \beta\gamma_6} \frac{R_k \parallel R_n}{R_r + r_6 + r_9}; \quad (7-16)$$

$$K_i = - \frac{\beta}{1 + \beta\gamma_6} \frac{R_r}{R_r + r_6 + r_9} \frac{R_k}{R_k + R_n}. \quad (7-17)$$

Смысл этих формул следующий. Если обратной связи нет, то ток эквивалентного генератора βI_6 , а следовательно, и выходные ток и напряжение определяются входным током I_6' (рис. 7-3). При наличии же обратной связи на ток I_6' накладывается ток обратной связи $(\beta I_6)\gamma_6$:

$$I_6 = I_6' - (\beta I_6) \gamma_6.$$

Отсюда результирующий ток

$$I_6 = \frac{I_6'}{1 + \beta\gamma_6}. \quad (7-18)$$

¹ Коэффициент K_u получается из соотношения $U_{вых} = -(\beta I_6') (R_k \parallel R_n)$, где $I_6' = E_r / (R_r + r_6 + r_9)$, а коэффициент K_i после представления источника сигнала в виде генератора тока I_r , зашунтированного резистором R_r , — из соотношения $I_{вых} = -(\beta I_6') [R_k / (R_k + R_n)]$, где $I_6' = I_r [R_r / (R_r + r_6 + r_9)]$.

Как видим, ток базы уменьшается в $1 + \beta \gamma_6$ раз по сравнению со случаем отсутствия обратной связи. Соответственно уменьшаются выходные ток и напряжение, а значит, и коэффициент усиления.

Сравнивая выражения (7-16) и (7-17) с (7-13) и (7-14), легко заметить, что влияние обратной связи равносильно уменьшению коэффициента передачи β в $1 + \beta \gamma_6$ раз. Эквивалентный коэффициент передачи $\beta/(1 + \beta \gamma_6)$, учитывающий внутреннюю обратную связь, является тем параметром, который при анализе статического режима

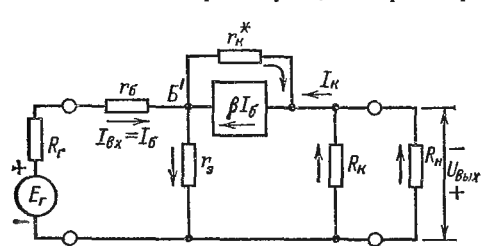


Рис. 7-4. Полная эквивалентная схема каскада в области средних частот.

$R_{вх}$ нужно полагать $R_r = 0$ (т. е. считать э. д. с. E_r приложенной непосредственно к входным зажимам: $U_{вх} = E_r$).

Из выражений (7-15)—(7-17) следует, что обратная связь является отрицательной (см. § 8-3), так как она увеличивает входное сопротивление и уменьшает коэффициенты усиления (см. § 8-2).

Полный анализ. Главной причиной, ограничивающей коэффициент усиления K_u при высокоомной нагрузке, является сопротивление r_k^* (рис. 7-4), которое не было учтено при выводе формулы (7-4а). В ряде случаев оно оказывает влияние и на другие параметры каскада.

Учесть наличие сопротивления r_k^* можно разными способами, в том числе заново анализируя схему методом контурных токов. Однако более наглядным и удобным представляется следующий путь. Введем коэффициент токораспределения в коллекторной цепи:

$$\gamma_k^* = \frac{I_k}{\beta I_6} = \frac{r_k^*}{r_k^* + [R_k \parallel R_n + r_e \parallel (r_6 + R_r)]}. \quad (7-19a)$$

Этот коэффициент показывает, какая доля тока βI_6 поступает во внешнюю цепь коллектора (остальная часть тока ответвляется в резистор r_k^*)¹.

Во всех практических случаях $R_k \parallel R_n \gg r_e$, поэтому (7-19а) упрощается:

$$\gamma_k^* \approx \frac{r_k^*}{r_k^* + R_k \parallel R_n}. \quad (7-19б)$$

¹ Выражение (7-19а) имеет ту же структуру, что и (7-10), так как в обоих случаях ток распределяется по двум ветвям. В схеме на рис. 7-4 одна ветвь содержит резистор r_k^* , а вторая — сложный резистор, эквивалентное сопротивление которого в выражении (7-19а) помещено в квадратные скобки.

каскада мы называли коэффициентом неустойчивости [см. (6-7)].

Тот факт, что в выражении для входного сопротивления стоит коэффициент обратной связи γ_{60} , а не γ_6 , объясняется тем, что параметр $R_{вх}$ определяется отношением $U_{вх}/I_6$, а не E_r/I_6 . Следовательно, для нахождения

С помощью коэффициента γ_k^* легко внести поправки в выражения (7-15)—(7-17). А именно, из (7-19а) следует соотношение $I_k = \gamma_k^* (\beta I_6)$, которое можно переписать следующим образом:

$$I_k = (\beta \gamma_k^*) I_6.$$

Как видим, уменьшение коллекторного тока в реальной схеме (рис. 7-4) по сравнению с идеализированной (рис. 7-2) равносильно уменьшению коэффициента β .

Таким образом, с формальной точки зрения учет сопротивления r_k^* означает замену коэффициента усиления β эквивалентным коэффициентом усиления

$$\beta_{oe} = \beta \gamma_k^* \approx \beta \frac{r_k^*}{r_k^* + R_k \parallel R_H}, \quad (7-20)$$

который следует использовать вместо β во всех ранее выведенных формулах. При этом, поскольку $\beta_{oe} < \beta$, входное сопротивление и коэффициенты усиления каскада с учетом r_k^* уменьшаются. Количественно эти поправки тем более существенны, чем больше суммарное сопротивление $R_k \parallel R_H$.

Теперь можно решить вопрос о максимальном коэффициенте усиления каскада. Заменяя в формуле (7-16) β на β_{oe} , полагая $R_k \parallel R_H = \infty$ и учитывая (4-90), нетрудно получить предельное значение

$$K_{u\infty} = \frac{r_k}{R_r + r_6 + r_9}, \quad (7-21a)$$

которое отнюдь не равно бесконечности, как следует из упрощенной теории. Однако формула (7-21a) дает такие большие значения $K_{u\infty}$, при которых не выполняется условие (7-1). Поэтому, чтобы оценить реальное значение максимального коэффициента усиления, нужно учесть обратную связь по напряжению с помощью генератора $\mu_{\text{ЭК}} U_k$ (см. рис. 4-24). Для этого достаточно разделить коэффициент $K_{u\infty}$ на двучлен $1 + K_{u\infty} \mu_{\text{ЭК}}$, где $\mu_{\text{ЭК}}$ играет роль коэффициента обратной связи β в ламповых схемах [62]. Тогда, используя соотношение (4-26) при $\gamma = 1$, получаем:

$$K_{u\text{ макс}} = \frac{r_k}{R_r + r_6 + r_9 \left(1 + \frac{\beta}{2}\right)}. \quad (7-21б)$$

При большом значении R_r разница между формулами (7-21a) и (7-21б) невелика, но в предельном случае, когда $R_r = 0$, разница получается в несколько раз. Например, при $\beta = 40$; $r_k = 1$ МОм; $r_6 = 100$ Ом; $r_9 = 25$ Ом и $R_r = 0$ формула (7-21б) дает $K_{u\text{ макс}} \approx 1500$, тогда как из формулы (7-21a) следует $K_{u\infty} = 8000$. Напомним, что в обоих случаях предполагалось неравенство $R_k \parallel R_H \gg r_k^*$, которое

даже при холостом ходе ($R_n = \infty$) практически нельзя выполнить в связи с трудностями установления режима транзистора ¹.

Заметим, что параметры $K_{u\infty}$ и $K_{u\max}$ являются аналогами «статического» коэффициента усиления μ в лампах и МДП транзисторах. Соответственно эти параметры можно найти из известного соотношения $\mu = SR_i$, где S — крутизна и R_i — внутреннее (выходное) сопротивление транзистора. Так же, как и в случае пентодов, коэффициент μ в случае биполярных транзисторов практически не используется, хотя и дает представление о предельном коэффициенте усиления каскада.

Заменяя β эквивалентным коэффициентом β_{oe} и сохраняя структуру упрощенной схемы (рис. 7-2), мы тем самым оставляем прежнее значение выходного сопротивления, определяемое формулой (7-3). Между тем из рис. 7-4 видно, что выходное сопротивление должно зависеть от значения r_k^* . Поэтому найдем значение $R_{\text{вых}}$ из рис. 7-4 обычным путем, т. е. задавая некоторую э. д. с. E на выходе при отключенной нагрузке R_n и при $E_r = 0$. Тогда ток I , обусловленный э. д. с. E , будет состоять из трех компонентов: I_{R_k} , $I_{r_k^*}$ и βI_b . Первый из них равен E/R_k . Второй компонент положим равным E/r_k^* , так как точку B' можно считать заземленной ². Наконец третий компонент (βI_b) найдем из соотношения

$$I_b = -I_{r_k^*}\gamma_6 - \beta I_b\gamma_6,$$

основанного на принципе суперпозиции ³.

Суммируя все три компонента и деля э. д. с. E на эту сумму, после некоторых преобразований получим:

$$R_{\text{вых}} = R_k \parallel [r_k^* (1 + \beta\gamma_6)]. \quad (7-22)$$

Здесь R_k является в н е ш н е й составляющей, а собственно выходное сопротивление со стороны коллектора, обусловленное величиной r_k^* , имеет вид:

$$R_{\text{вых}} = r_k^* (1 + \beta\gamma_6). \quad (7-23)$$

В обеих полученных формулах коэффициент β имеет обычный смысл и не должен заменяться на β_{oe} . Выходное сопротивление увеличивается с увеличением γ_6 .

¹ В первую очередь речь идет о падении напряжения на резисторе R_k от постоянной составляющей коллекторного тока. Например, при токе $I_k = 1$ мА сопротивление r_k^* обычно составляет не менее 10 кОм. Если принять $R_k = 5r_k^*$, то напряжение $I_k R_k$ превысит 50 В, а напряжение питания E_k будет еще больше.

² В самом деле, если бы генератор βI_b отсутствовал, то напряжение на эмиттерном переходе составляло бы $U_{r_9} \approx E \frac{r_9}{r_k^*} \ll E$; при наличии генератора βI_b напряжение U_{r_9} оказывается еще меньше.

³ В данном случае используется суперпозиция (сложение) двух составляющих базового тока I_b' и I_b'' , первая из которых обусловлена э. д. с. E и сопротивлением r_k^* (при условии $\beta = 0$ и $r_9 = 0$), а вторая — генератором тока βI_b (при условии $E = 0$).

Максимальное значение $R_{\text{вых}} = r_k^* (1 + \beta \gamma_{60})$ получается при $R_r = 0$, когда источником сигнала является генератор э. д. с. Минимальное значение $R_{\text{вых}} = r_k^*$ получается при $R_r = \infty$, когда источником сигнала является генератор тока.

Еще один фактор, который не учитывается при использовании эквивалентного коэффициента β_{oe} и упрощенной эквивалентной схемы (рис. 7-2) — это непосредственное, «сквозное» прохождение части входного сигнала на выход через делитель напряжения $r_k^* - R_k \parallel R_H$, подключенный параллельно сопротивлению r_g . Из рис. 7-4, полагая $\beta = 0$ и учитывая соотношение $r_g \ll r_k^*$, нетрудно получить коэффициент сквозной передачи в виде

$$K_{\text{скв}} = \frac{(U_{\text{вых}})_{\text{скв}}}{E_r} \approx \gamma_6 (1 - \gamma_k^*).$$

В отличие от нормального коэффициента усиления K_u коэффициент $K_{\text{скв}}$ имеет положительную полярность, так как при непосредственном прохождении сигнала его полярность не меняется. Обычно $K_{\text{скв}} \ll 1$, тогда как $K_u \gg 1$; следовательно, роль сквозного прохождения незначительна, во всяком случае в области средних частот.

7-3. КАСКАД В ОБЛАСТИ БОЛЬШИХ ВРЕМЕН И НИЗШИХ ЧАСТОТ

Эквивалентная схема каскада в области больших времен, в которой существенное значение имеют емкости C_1 , C_2 и C_3 , показана на рис. 7-5. Здесь для простоты не учтено сопротивление r_k^* , так как указанные емкости выбираются всегда со значительным запасом, который делает нецелесообразным учет тех поправок, которые обусловлены величиной r_k^* . Рассмотрим влияние каждой из этих емкостей поочередно, полагая остальные две емкости бесконечно большими (т. е. «закороченными»).

Влияние переходных емкостей. Положим сначала $C_2 = \infty$, $C_3 = \infty$ и выясним роль C_1 . В первый момент после поступления ступенчатого сигнала наличие этой емкости не имеет значения и входной ток, а значит, и выходные величины будут такими же, как в области средних частот¹. В дальнейшем емкость C_1 заряжается и входной ток, а значит, и выходные величины уменьшаются. Получается спад вершины выходного импульса. Для количественного анализа следует добавить к R_r операторное сопротивление $1/(sC_1)$. Полагая теперь $C_1 = \infty$, $C_r = \infty$ и рассматривая влияние C_2 , приходим к выводу, что и в данном случае в первый момент схема

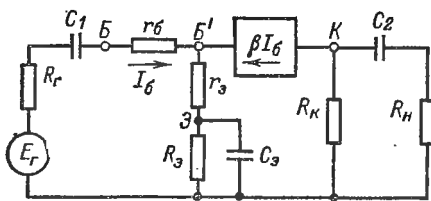


Рис. 7-5. Эквивалентная схема каскада в области больших времен (низших частот).

¹ Для ступенчатого сигнала емкость в первый момент можно рассматривать как короткозамкнутый участок.

работает так же, как в области средних частот. В дальнейшем, по мере заряда емкости C_2 , входной ток не меняется, а ток и напряжение нагрузки уменьшаются. Для количественного анализа следует добавить к R_n операторное сопротивление $1/(sC_2)$.

В обоих случаях из выражения (7-4а) нетрудно получить изображение

$$K_u(s) = K_{u0} \frac{s\tau_n}{1 + s\tau_n}, \quad (7-24)$$

в котором K_{u0} — коэффициент усиления в области средних частот, а постоянная времени τ_n имеет вид:

$$\tau_{n1} = C_1 (R_r + R_{вх}) \quad (7-25a)$$

для первой емкости и

$$\tau_{n2} = C_2 (R_k + R_n) \quad (7-25b)$$

для второй емкости.

Оригиналом изображения (7-24) является обычная экспоненциально спадающая функция (рис. 7-6):

$$K_u(t) = K_{u0} e^{-\frac{t}{\tau_n}}. \quad (7-26)$$

Относительный спад вершины при достаточно коротком импульсе ($t_n \ll \tau_n$) выражается известной формулой [62]:

$$\delta = \frac{\Delta U_{ввх}}{U_{ввх}} \approx \frac{t_n}{\tau_n} \quad (7-27)$$

(см. рис. 7-6).

В случае транзисторных каскадов получить малый спад δ часто нелегко, поскольку постоянные времени τ_{n1} и τ_{n2} можно увеличивать только путем увеличения емкостей. Последние должны быть весьма большими, так как сопротивления R_r , $R_{вх}$, R_k и R_n обычно не превышают нескольких килоом. Как правило, конденсаторы C_1 и C_2 — электролитические, с емкостями в десятки, а иногда и сотни микрофард.

Частотные характеристики, обусловленные переходными конденсаторами, получаются из выражения (7-24) заменой оператора s на $j\omega$:

$$\dot{K}_u = \frac{K_{u0}}{1 - j \frac{\omega_n}{\omega}}. \quad (7-28)$$

Нижние граничные частоты ω_{n1} и ω_{n2} определяются как величины, обратные постоянным времени τ_{n1} и τ_{n2} [см. (7-25)]. Частотная характеристика имеет такой же вид, как и в ламповом каскаде (рис. 7-7).

Влияние блокирующей емкости в цепи эмиттера. Положим $C_1 = \infty$, $C_2 = \infty$ и выясним роль емкости C_e в переходном процессе.

В первый момент после поступления сигнала емкость C_3 заряжается резистор R_3 , т. е. начальные значения токов и напряжений такие же, как на средних частотах. По мере заряда конденсатора C_3 уменьшается эмиттерный ток; вместе с ним уменьшаются ток базы и выходные величины. В пределе, когда конденсатор C_3 зарядится полностью и ток через него сделается равным нулю, сопротивление в эмиттерной цепи станет равным $R_3 + r_e$ вместо начального значения r_e . Очевидно, что ток базы и выходные величины

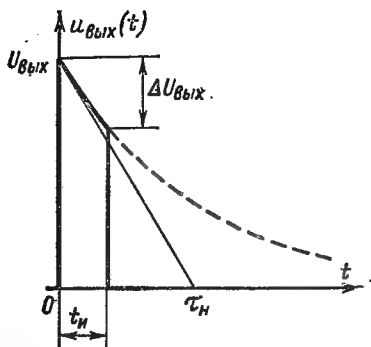


Рис. 7-6. Переходная характеристика в области больших времен (штриховая линия) и прямоугольный импульс с искаженной вершиной (жирная линия).

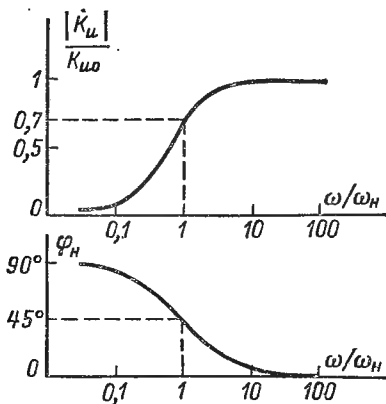


Рис. 7-7. Амплитудно- и фазо-частотные характеристики в области низких частот (сдвиг фазы на 180° в области средних частот не учитывается).

сильно уменьшаются за время переходного процесса, но их конечные значения будут все же отличны от нуля. В этом заключается принципиальная особенность влияния емкости C_3 по сравнению с влиянием емкостей C_1 и C_2 . Постоянная времени переходного процесса равна произведению емкости C_3 на суммарное шунтирующее ее сопротивление. Последнее представляет собой параллельное соединение внешнего сопротивления R_3 и внутреннего выходного сопротивления каскада со стороны эмиттера.

Для количественного анализа можно использовать формулу (7-4а), добавив к сопротивлению r_e (входящему в $R_{вх}$) операторное сопротивление

$$Z_3(s) = \frac{R_3}{1 + s\tau'_{н.3}},$$

где $\tau'_{н.3} = C_3 R_3$. После этого легко привести изображение коэффициента усиления к следующему виду:

$$K_u(s) = \frac{K_{u0}}{1 + \frac{(1 + \beta) R_3 / (R_г + R_{вх})}{1 + s\tau'_{н.3}}}. \quad (7-29a)$$

Отсюда видно, что в предделе, при $t = \infty$ (т. е. при $s = 0$), коэффициент усиления будет гораздо меньше его начального значения. Так, при $R_r = 0$ и $r_6 \ll (1 + \beta) r_3$ получаем $K_u(\infty)/K_u(0) \approx \approx r_3/R_3$, где r_3 не превышает десятков ом, а R_3 составляет 1 кОм и больше.

Практически можно считать $K_u(\infty) = 0$ и в формуле (7-29а) пренебречь единицей по сравнению с членом $s\tau'_{н.э}$. Тогда изображение $K_u(s)$ будет иметь ту же форму, что и (7-24):

$$K_u(s) \approx K_{u0} \frac{s\tau_{н.э}}{1 + s\tau_{н.э}}, \quad (7-29б)$$

где постоянная времени равна:

$$\tau_{н.э} = C_3 \frac{R_r + R_{вх}}{1 + \beta} = C_3 \left(r_3 + \frac{R_r + r_6}{1 + \beta} \right). \quad (7-30а)$$

Как видим, в таком приближении переходная характеристика является спадающей экспонентой (7-26), а ее постоянная времени не зависит от сопротивления R_3 . Это значит, что последнее значительно больше, чем выходное сопротивление со стороны эмиттера, с которым оно включено параллельно. Если положить $R_r + r_6 \ll \ll (1 + \beta) r_3$, что иногда имеет место, то постоянная времени запишется элементарно просто:

$$\tau_{н.э} \approx C_3 r_3. \quad (7-30б)$$

Сравнивая это выражение с выражениями (7-25), легко сделать вывод, что емкость C_3 при прочих равных условиях должна быть значительно больше, чем переходные емкости, так как $r_3 \ll R_{вх}$ и $r_3 \ll R_k$. Типичными значениями блокирующей емкости являются сотни микрофард. Искажения вершины коротких импульсов рассчитываются по формуле (7-27).

Частотные характеристики, обусловленные емкостью C_3 , получаются заменой оператора s на $j\omega$ в (7-29). Если исходить из формулы (7-29б), приходим к обычному выражению (7-28), в котором нижняя граничная частота $\omega_{н.э}$ есть величина, обратная постоянным времени (7-30). Если исходить из полной формулы (7-29а), то при $\omega = 0$ получается к о н е ч н о е значение K_u , что соответствует структуре эквивалентной схемы.

Совместное влияние емкостей. Если одна из постоянных времени ($\tau_{н1}$, $\tau_{н2}$ или $\tau_{н.э}$) много меньше двух остальных, то именно она определяет переходную и частотную характеристики. В этом случае в формулы (7-26) и (7-28) следует подставлять наименьшее значение τ_n или наибольшее значение ω_n . Если же постоянные времени сравнимы, то анализ переходных и частотных характеристик усложняется. Однако в таких случаях главный интерес обычно представляют не сами характеристики, а результирующий спад вершины короткого импульса и результирующая граничная частота.

Для оценки этих величин можно воспользоваться следующими выводами из теории ламповых каскадов [62].

Результирующий спад вершины определяется формулой (7-27), если заменить в ней постоянную времени τ_n так называемым *временем спада вершины*:

$$\delta \approx \frac{t_n}{t_{\text{сп}}}. \quad (7-31)$$

Время спада, называемое иногда (менее точно) эквивалентной постоянной времени, выражается в следующем виде:

$$t_{\text{сп}} = \left(\sum_{i=1}^m \frac{1}{\tau_{ni}} \right)^{-1}, \quad (7-32a)$$

где τ_{ni} — постоянная времени i -й цепочки.

В распространенном случае, когда все m постоянные времени одинаковы, получаем:

$$t_{\text{сп}} = \frac{\tau_n}{m}. \quad (7-32b)$$

Значит, по заданным значениям δ и t_n нужно найти необходимое время $t_{\text{сп}}$, а затем рассчитать каждую емкость из соотношения $\tau_n = m t_{\text{сп}}$. В каскаде на рис. 7-1, очевидно, $m = 3$.

Результирующая граничная частота ω_{n0} на уровне 0,7 не имеет достаточно простой связи с граничными частотами отдельных цепочек ω_{ni} . Однако весьма приближенно ее можно оценить, заменив в формулах (7-32) величины τ_{ni} на $1/\omega_{ni}$ (что вполне строго) и величину $t_{\text{сп}}$ на $1/\omega_{n0}$ (что и является причиной неточности). Тогда в общем случае

$$\omega_{n0} \approx \sum_{i=1}^m \omega_{ni}, \quad (7-33a)$$

а в частном случае, когда все ω_{ni} одинаковы,

$$\omega_{n0} \approx m \omega_n. \quad (7-33b)$$

Соотношения (7-33) гораздо более точны тогда, когда граничные частоты определяются на уровне более высоком, чем 0,7, например на уровне 0,9.

Коррекция искажений вершины. Для уменьшения спада вершины можно применять RC-фильтр в коллекторной цепи (рис. 7-8). Этот способ хорошо известен из ламповой техники (см., например, [62], § 9-1). Поэтому, не проводя детального анализа, укажем лишь его результаты. Изображение для коэффициента усиления

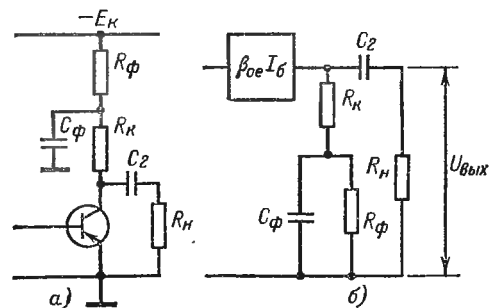


Рис. 7-8. Принципиальная (а) и эквивалентная (б) схемы каскада с коллекторным фильтром для коррекции искажений вершины импульса.

с учетом корректирующего фильтра имеет вид:

$$K_{\text{н}}(s) = K_{\text{н}0} \frac{1 + \frac{\alpha_{\phi}}{1 + s\tau_{\phi}}}{1 + \frac{1}{s\tau_{\text{н}}} + \frac{\alpha'_{\phi}}{1 + s\tau_{\phi}}},$$

где $\tau_{\phi} = C_{\phi}R_{\phi}$; $\tau_{\text{н}} = C_2R_{\text{н}}$; $\alpha_{\phi} = R_{\phi}/R_{\text{к}}$; $\alpha'_{\phi} = R_{\phi}/(R_{\text{н}} + R_{\text{к}})$.

Условие коррекции вершины, при котором отсутствует линейный спад в начале выходного импульса, оказывается таким же, как и в случае лампового каскада:

$$C_{\phi}R_{\text{к}} = C_2R_{\text{н}}. \quad (7-34)$$

Отсюда определяют необходимую емкость фильтра C_{ϕ} . При выполнении условия (7-34) спад вершины будет квадратичным, а значение относительного спада в конце импульса будет выражаться формулой

$$\delta = \frac{t_{\text{н}}^2}{2\alpha_{\phi}\tau_{\text{н}}^2},$$

где $t_{\text{н}}$ — длительность импульса.

При наличии нескольких низкочастотных RC -цепочек (см. рис. 7-1) вместо постоянной времени $C_2R_{\text{н}}$ в правой части (7-34) следует использовать время спада (7-32).

7-4. КАСКАД В ОБЛАСТИ МАЛЫХ ВРЕМЕН И ВЫСШИХ ЧАСТОТ

Эквивалентная схема каскада в этой области показана на рис. 7-9. Здесь в отличие от схемы на рис. 7-2 учтена емкость коллекторного перехода $C_{\text{к}}^*$, а коэффициент передачи β считается операторной величиной. Поскольку параметры $C_{\text{к}}^*$ и $r_{\text{к}}^*$ являются функциями β , они тоже являются операторными величинами, что отражено точками над их обозначениями. Как обычно, при импульсном анализе мы будем пользоваться оператором s , а при гармоническом — оператором $j\omega$. Емкость эмиттерного перехода на рис. 7-9 не учитывается¹. Нагрузку будем считать активной.

Переходные характеристики. Начнем с качественного рассмотрения переходных процессов, причем сначала пренебрежем влиянием цепочки $C_{\text{к}}^*$, $r_{\text{к}}^*$.

Если на входе задана ступенька тока I_6 , то в первый момент, когда $\beta = 0$, входное сопротивление выражается формулой (7-12), а выходные величины равны нулю. В дальнейшем β экспоненциально увеличивается с постоянной времени τ_{β} и величины $R_{\text{вх}}$, $I_{\text{к}}$ и $U_{\text{вых}}$ возрастают вплоть до установившихся значений, соответствующих области средних частот.

Пусть теперь задана ступенька э. д. с. $E_{\text{г}}$ при конечном сопротивлении $R_{\text{г}}$. Тогда в первый момент ток I_6' будет определяться суммой сопротивлений $R_{\text{г}} + r_6 + r_3$, а входные величины (как и в случае заданной ступеньки тока) будут равны нулю. По мере нарастания коэффициента β увеличивается ток $I_{\text{к}}$ и часть

¹ В случае дрейфовых транзисторов повышенное значение C_3 компенсируется увеличением рабочего тока, т. е. уменьшением r_3 (см. § 4-12). В тех случаях, когда емкость C_3 существенна (т. е. постоянная времени τ_3 сравнима с τ_{α}), ее можно учесть путем изменения величины τ_{α} [см. примечания к (4-57) — (4-59)].

его ответвляется в базовую цепь каскада на вст реч у исходному току I_6^* . Соответствующее уменьшение тока базы во время переходного процесса способствует более быстрому завершению процесса. Это равносильно уменьшению постоянной времени по сравнению со значением τ_β , которая характеризовала переходный процесс при постоянном токе базы. Такой же вывод следует из общих свойств отрицательной обратной связи, которая имеет место при $R_r \neq \infty$ (см. раздел «Внутренняя обратная связь» в § 7-2).

Рассмотрим влияние цепочки C_k^* , r_k^* .

В первый момент после поступления ступеньки E_r , когда $\beta = 0$, параметры C_k^* и r_k^* равны соответственно C_k и r_k , т. е. гораздо слабее шунтируют генератор βI_6 , чем в установившемся режиме. Однако в этот первый момент сопротивление $R_k \parallel R_n$ оказывается присоединенным параллельно r_3 (через емкость C_k) и, следовательно, на выходе имеет место небольшой начальный скачок напряжения за счет частичного непосредственного прохождения сигнала. Полярность этого начального скачка та же, что и у сигнала E_r , в отличие от последующих этапов, когда полярность меняется на обратную, свойственную установившемуся режиму. По мере возрастания β емкость C_k^* увеличивается, а сопротивление r_k^* уменьшается. Значит, все большая доля тока βI_6 ответвляется в цепь Z_k^* , что ослабляет обратную связь и затягивает переходный процесс.

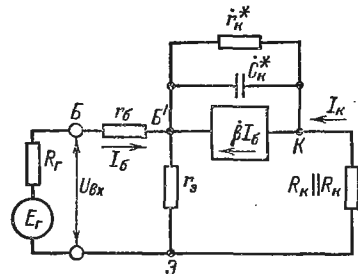


Рис. 7-9. Эквивалентная схема каскада в области малых времен (высших частот).

Теперь подтвердим все сказанное количественным анализом. Для этого используем формулы, полученные в § 7-2, заменив в них соответствующие действительные величины операторными, импеданс $Z_k^*(s)$ заимствуем из формулы (4-49б), а для коэффициента $\beta(s)$ примем выражение (4-80а). Подставляя эти значения в определение (7-20), после некоторых преобразований получаем изображение эквивалентного коэффициента $\beta_{oe}(s)$ в следующем виде:

$$\beta_{oe}(s) = \frac{\beta_{oe}}{1 + s\tau_{oe}}. \quad (7-35)$$

Эквивалентный коэффициент передачи β_{oe} в правой части определяется формулой (7-20) для средних частот, а эквивалентная постоянная времени τ_{oe} — соотношением

$$\tau_{oe} = \tau'_\beta + \tau'_k, \quad (7-36)$$

где

$$\tau'_\beta = \gamma_k^* \tau_\beta; \quad \tau'_k = (1 - \gamma_k^*) \tau_k,$$

а γ_k^* — коэффициент токораспределения (7-19б) ¹.

¹ Постоянную времени τ'_k можно представить в такой наглядной форме:

$$\tau'_k = C_k^* (r_k^* \parallel R_k \parallel R_n),$$

где емкость C_k^* умножена на суммарное сопротивление всех ветвей, по которым разветвляется ток βI_6 .

Очевидно, что величина τ_{oe} зависит от γ_K^* линейно, меняясь от значения τ_β при $\gamma_K^* = 1$ до значения τ_K при $\gamma_K^* = 0$. В зависимости от того, какая из постоянных времени (τ_β или τ_K) больше, эквивалентная постоянная времени либо уменьшается, либо возрастает с увеличением суммарной нагрузки $R_K \parallel R_H$, определяющей величину γ_K^* . Поскольку сопротивление r_K^* , а вместе с ним и τ_K обратно пропорциональны току [см. (4-24)], соотношение величин τ_β и τ_K не однозначно для данного транзистора, а зависит от режима. Следовательно, от режима зависят как величина τ_{oe} , так и характер функции $\tau_{oe}(\gamma_K^*)$.

С помощью изображения $\beta_{oe}(s)$ нетрудно найти изображения входного импеданса и коэффициента усиления. Подставляя $\beta_{oe}(s)$ в формулу (7-15), получаем:

$$Z_{вх}(s) = (r_6 + r_3) \left(1 + \frac{\beta_{oe} \gamma_{60}}{1 + s \tau_{oe}} \right). \quad (7-37)$$

Этому изображению соответствует оригинал

$$z_{вх}(t) = (r_6 + r_3) \left[1 + \beta_{oe} \gamma_{60} \left(1 - e^{-\frac{t}{\tau_{oe}}} \right) \right], \quad (7-38)$$

который иллюстрируется кривыми на рис. 7-10. Отсюда видно, что входное сопротивление во время переходного процесса увеличи-

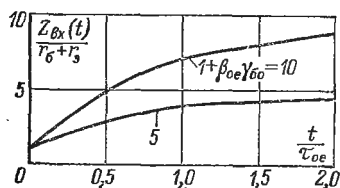


Рис. 7-10. Переходные характеристики входного сопротивления.

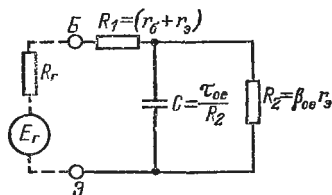


Рис. 7-11. Эквивалентная схема входа каскада ОЭ в области малых времен.

вается: Это обстоятельство можно отразить эквивалентной схемой, показанной на рис. 7-11. Из схемы ясно, что *входной импеданс имеет емкостный характер*. Дополнительное подтверждение такого вывода дают кривые входного напряжения и входного тока, построенные с помощью функции (7-38) и показанные на рис. 7-12.

Подставляя $\beta_{oe}(s)$ в формулу (7-16), получаем изображение коэффициента усиления напряжения:

$$K_u(s) = \frac{K_{u0}}{1 + s \frac{\tau_{oe}}{1 + \beta_{oe} \gamma_6}}. \quad (7-39)$$

Оригиналом этого изображения является простая экспоненциальная функция

$$k_u(t) = K_{u0} \left(1 - e^{-\frac{t}{\tau_{в}}} \right), \quad (7-40)$$

где τ_b — постоянная времени в области высших частот:

$$\tau_b = \frac{\tau_{oe}}{1 + \beta_{oe} \gamma_6}. \quad (7-41a)$$

Во многих схемах достаточно сильно выполняется неравенство $\beta_{oe} \gamma_6 > 1$, так что в знаменателе (7-41a) можно пренебречь единицей. Тогда, используя выражения (7-36), (7-19б) и (7-10), нетрудно получить:

$$\tau_b \approx \frac{1}{\gamma_6} (\tau_a + C_k \cdot R_k \parallel R_H). \quad (7-41б)$$

В этом случае инерционность каскада определяется параметрами, свойственными схеме ОБ.

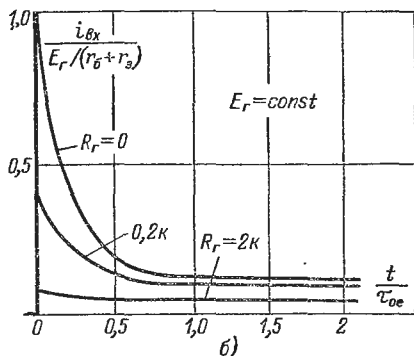
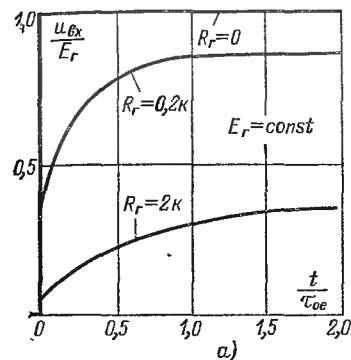


Рис. 7-12. Переходные характеристики входного напряжения (а) и входного тока (б) при неизменном значении E_r и меняющемся R_r .

Функция $k_u(t)$ иллюстрируется кривыми на рис. 7-13. Постоянная времени τ_b существенно зависит от сопротивления источника сигнала R_r через коэффициент γ_6 и от сопротивления суммарной нагрузки $R_k \parallel R_H$ через коэффициент γ_k^* . Чем меньше сопротивление R_r , тем меньше постоянная времени τ_b и тем короче фронт. Одновременно согласно (7-4a) увеличивается коэффициент усиления K_{u0} . Что касается влияния суммарной нагрузки, то оно будет разным в зависимости от соотношения значений τ_b и τ_k . Если $\tau_b < \tau_k$, как обычно и бывает, то постоянная времени τ_b возрастает с ростом нагрузочного сопротивления. Время нарастания фронта на уровне 0,1—0,9 выражается соотношением

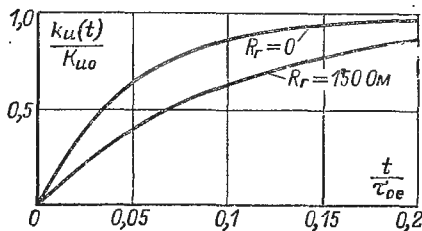


Рис. 7-13. Переходные характеристики коэффициента усиления.

$$t_n = 2,2 \tau_b = 2,2 \frac{\tau'_b + \tau'_k}{1 + \beta_{oe} \gamma_6}. \quad (7-42a)$$

При $R_r = 0$ и $R_k \parallel R_n \ll r_k^*$ время нарастания минимально и определяется выражением

$$t_{н. мин} \approx 2,2 \frac{\tau_\beta}{1 + \beta \gamma_{60}}. \quad (7-426)$$

Например, при $\tau_\beta = 2$ мкс; $\beta = 50$ и $\gamma_{60} = 0,2$ получим $t_{н. мин} = 0,4$ мкс.

В заключение заметим, что если в основу анализа положить не упрощенное выражение для $Z_k^*(s)$ в виде (4-94б), а более точное выражение (4-94а), то в изображениях (7-37) и (7-39) появятся члены с s^2 . Аперриодический характер функций $z_{вх}(t)$ и $k_u(t)$ при этом сохранится, но последняя приобретет структуру, свойственную двухкаскадному усилителю с сильно различающимися постоянными времени. Время нарастания при этом практически не изменится, но появится небольшая задержка фронта на время порядка τ_α .

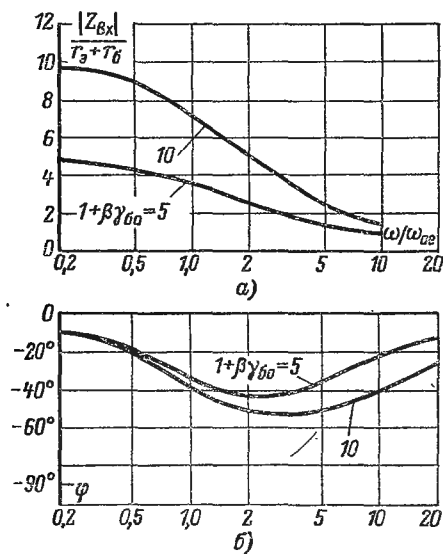


Рис. 7-14. Частотные характеристики входного сопротивления.

а — амплитудно-частотные; б — фазо-частотные.

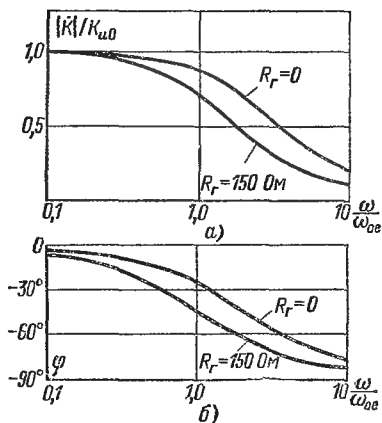


Рис. 7-15. Частотные характеристики коэффициента усиления.

а — амплитудно-частотные; б — фазо-частотные.

Частотные характеристики. Заменяв оператор s на $j\omega$ в формуле (7-37), получим частотную характеристику входного сопротивления:

$$Z_{вх} = (r_6 + r_9) \left(1 + \frac{\beta_{0e} \gamma_{60}}{1 + j\omega/\omega_{0e}} \right), \quad (7-43)$$

где частота ω_{0e} есть величина, обратная постоянной времени τ_{0e} (7-36).

Амплитудно- и фазо-частотные характеристики показаны на рис. 7-14. Как видим, модуль $Z_{вх}$ с увеличением частоты уменьшается и стремится к предельной величине $r_6 + r_9$. Фаза $Z_{вх}$ отрицательна; она сначала возрастает (по модулю), достигает максимума, а затем

асимптотически уменьшается до нуля. Анализируя функцию $\varphi(\omega)$, нетрудно показать, что она имеет максимум при частоте ¹

$$\omega_{\varphi m} = \omega_{oe} \sqrt{1 + \beta_{oe} \gamma_6}$$

и что максимум тем ближе к 90° , чем больше $\sqrt{1 + \beta_{oe} \gamma_6}$.

Из формулы (7-39) заменой оператора получаем частотную характеристику коэффициента усиления:

$$\dot{K}_u = \frac{K_{uo}}{1 + j \frac{\omega}{\omega_B}}, \quad (7-44)$$

где верхняя граничная частота ω_B есть величина, обратная постоянной времени τ_B :

$$\omega_B = \frac{1 + \beta_{oe} \gamma_6}{\tau'_\beta + \tau'_K}. \quad (7-45)$$

Частотные характеристики K_u иллюстрируются кривыми на рис. 7-15.

Иногда представляет интерес частотная характеристика выходного сопротивления каскада. Подставив в (7-23) выражение для Z_K^* из (4-94б) и выражение для β из (4-80б), после небольших преобразований получим искомую характеристику в следующем виде:

$$Z_{\text{вых}} = R_{\text{вых}} \frac{1 + j \frac{\omega \tau_\beta}{1 + \beta \gamma_6}}{1 + j \omega \tau_K}. \quad (7-46)$$

То, что выходное сопротивление на высоких частотах имеет конечную (хотя и очень малую) величину, объясняется приближенностью формулы (4-94б). Амплитудно- и фазо-частотная характеристики $Z_{\text{вых}}$ в общем подобны характеристикам $Z_{\text{вх}}$ на рис. 7-14.

Максимальная (по модулю) фаза получается в данном случае при частоте

$$\omega_{\varphi m} = \sqrt{\frac{1 + \beta \gamma_6}{\tau_\beta \tau_K}}$$

и оказывается тем ближе к 90° , чем больше τ_K и чем меньше τ_β .

Добротность каскада. Из ламповой техники известно, что время нарастания фронта в каскаде прямо пропорционально, а верхняя граничная частота обратно пропорциональна коэффициенту усиления. Постоянной величиной, определяющей отношение K/t_n или произведение Kf_B , служит *добротность лампы* или, точнее, каскада ². Пренебрегая числовым множителем, который зависит от способа

¹ Следует напомнить, что фазо-частотная характеристика коэффициента β , вытекающая из (4-80б) и лежащая в основе фазо-частотных характеристик усилительного каскада, является неточной, особенно в области высоких частот [см. (4-83б)]. Поэтому количественные выводы относительно функции $\varphi(\omega)$ следует считать ориентировочными.

² Вместо термина «добротность» для величин $K\omega_B$ и Kf_B часто применяют термин *площадь усиления*.

определения добротности (K/t_n , K/τ_a , $K\omega_n$ или Kf_n), можно считать ее равной S/C , где S — крутизна лампы, а C — сумма входной и выходной емкостей. Добротность характеризует предельные возможности каскада, т. е. минимальное время нарастания (или максимальную частоту сигнала) при $K > 1$.

Постоянство добротности в ламповом каскаде обусловлено тем, что и коэффициент усиления и время нарастания пропорциональны одной и той же величине — суммарной анодной нагрузке $R_a \parallel R_n$. В транзисторном каскаде такая пропорциональность в общем случае не имеет места, поэтому добротность не обязательно является постоянной величиной. В этом заключается существенное отличие транзисторного каскада от лампового. Рассмотрим этот вопрос подробнее на примере того однокаскадного усилителя, который анализировался выше.

Примем за добротность отношение K_{uo}/t_n и обозначим ее через D . Тогда, разделив выражение (7-16) на (7-42а) и опустив знак минус, получим:

$$D = \frac{0,45\beta_{oe}}{\tau'_\beta + \tau'_k} \frac{R_k \parallel R_n}{R_r + r_6 + r_9}. \quad (7-47)$$

При малых значениях суммарной нагрузки $R_k \parallel R_n$, когда $\beta_{oe} \approx \beta$, $\tau'_\beta \approx \tau_\beta$ и $\tau'_k \ll \tau_\beta$, добротность растет пропорционально $R_k \parallel R_n$ и определяется постоянной времени τ_α :

$$D \approx \frac{0,45}{\tau_\alpha} \frac{R_k \parallel R_n}{R_r + r_6 + r_9}. \quad (7-48a)$$

При больших значениях суммарной нагрузки, когда $\tau'_k \gg \tau'_\beta$, добротность достигает максимальной величины и выражение (7-47) переходит в следующее:

$$D \approx \frac{0,45}{C_k (R_r + r_6 + r_9)}. \quad (7-48б)$$

В этом случае добротность, как и в ламповом каскаде, не зависит от сопротивления нагрузки¹.

Исключительно важным является то обстоятельство, что максимальная добротность (7-58б) зависит не только от параметров транзистора, но и от сопротивления источника сигнала R_r , причем увеличение R_r снижает добротность. Эта особенность в какой-то степени аналогична тому, что в ламповой технике добротность каскада определяется не только добротностью лампы,

¹ Условие $\tau'_k \gg \tau'_\beta$, использованное при выводе (7-48б), приводится к виду

$$R_k \parallel R_n \gg \frac{\tau_\alpha}{C_n}.$$

Выполнение такого неравенства часто затруднительно, а иногда невозможно, особенно в случае бездрейфовых транзисторов, для которых правая часть неравенства может составлять 5 кОм и более.

но и внешней монтажной емкостью. Однако роль R_r гораздо более существенна¹.

Полагая $R_r = 0$ в формуле (7-48б), получаем добротность самого транзистора, которая одновременно является предельной добротностью усилителя:

$$D_0 = \frac{0,45}{C_k(r_6 + r_9)}. \quad (7-49a)$$

Обычно $r_9 \ll r_6$ и можно приближенно считать:

$$D_0 \approx \frac{0,45}{C_k r_6}. \quad (7-49б)$$

Отсюда видно, что в данном случае добротность определяется не постоянной времени τ_a [см. (7-48а)], а постоянной времени $\tau_6 = C_k r_6$, значение которой уже подчеркивалось в § 4-5.

До сих пор анализировалась добротность самостоятельного усилительного каскада с произвольными значениями R_r и R_n . Если иметь в виду промежуточный каскад (см. сноску на с. 328), то в исходном выражении (7-47) следует положить $R_n = \infty$ [см. (7-57)] и $R_r = R_{\text{вых}} \approx R_k$ [см. (7-3)]. Тогда, раскрывая значения τ_β и τ'_k согласно (7-36) и (7-19б), а также учитывая соотношения (4-81а) и (4-92), можно привести добротность каскада к следующему общему виду:

$$D \approx \frac{0,45}{\tau_a + C_k R_k} \frac{R_k}{R_k + r_6 + r_9}. \quad (7-50a)$$

Легко убедиться что эта добротность имеет максимальное значение

$$D_m \approx \frac{0,45}{C_k(r_6 + r_9)} \frac{1}{1 + \sqrt{\tau_a / C_k(r_6 + r_9)}} \quad (7-50б)$$

при некотором оптимальном значении режимного сопротивления R_k :

$$R_{k, \text{опт}} = \sqrt{\frac{\tau_a}{C_k}(r_6 + r_9)}. \quad (7-51)$$

Такое значение рекомендуется использовать в многокаскадных импульсных усилителях [115]. Сравнивая выражения (7-50б) и (7-49б), видим, что у промежуточного каскада максимальная добротность меньше, чем у однокаскадного усилителя. Обе они зависят от эмиттерного сопротивления, т. е. от тока $I_{\text{э}}$.

Если каскад предназначен для усиления синусоидальных сигналов, то под добротностью удобнее понимать произведение $K_{\text{эф}}$. Тогда в приведенных выше выражениях нужно вместо коэффициента 0,45 использовать коэффициент 1,3.

¹ Например, даже при $R_r = 100 \text{ Ом}$ добротность D , как правило, в 2—3 раза меньше, чем при $R_r = 0$. При значениях $R_r \gg r_6 + r_9$ значение D уменьшается на порядок и более.

Помимо рассматриваемой добротности по напряжению можно использовать добротность по току, а в общем случае добротность по мощности [116].

Коррекция фронта. Из теории усилителей известно, что при соединении каскадов усиление растет быстрее, чем время нарастания, и добротность повышается. Поэтому с точки зрения переходных и частотных свойств выгоднее использовать два каскада вместо одного при том же коэффициенте усиления. Естественно что при этом каждый из каскадов двухкаскадного усилителя имеет меньшее усиление, чем однокаскадный усилитель. Усиление каскада выгодно уменьшать не путем уменьшения сопротивлений в коллекторной

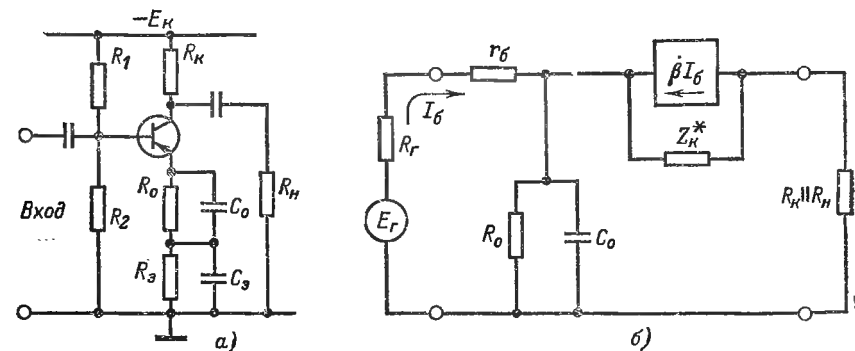


Рис. 7-16. Принципиальная (а) и эквивалентная (б) схемы каскада с обратной связью и коррекцией фронта.

цепи, а с помощью отрицательной обратной связи, вводя последовательное сопротивление R_0 в цепь эмиттера (рис. 7-16, а). Такая обратная связь повышает стабильность, линейность и входное сопротивление каскада. Однако одновременно снижается его добротность, так как к сопротивлению r_s в выражении (7-47) добавляется сопротивление R_0 , имеющее обычно гораздо большую величину. Снижение добротности каскада означает, что выигрыш от использования двух каскадов вместо одного уменьшается. Для восстановления добротности каскада сопротивление R_0 шунтируют небольшой емкостью C_0 , которую называют *корректирующей*. Идея такой коррекции заключается в том, чтобы «закоротить» R_0 на первом этапе переходного процесса. В результате будут сильно увеличены начальный ток базы и начальная скорость нарастания тока эквивалентного генератора βI_b . Соответственно сократится время нарастания фронта.

Поскольку R_0 вводится специально для создания отрицательной обратной связи, естественно считать $R_0 \gg r_s$. Для упрощения анализа положим $r_s = 0$. Тогда эквивалентная схема каскада будет такой, как показано на рис. 7-16, б. В области средних частот, когда емкостью C_0 можно пренебречь, коэффициент усиления будет

выражаться формулой (7-16), в которой r_0 следует заменить сопоставлением R_0 :

$$K_{u0} = \frac{\beta_{oe}(R_K \parallel R_H)}{(1 + \beta_{oe}\gamma_6)(R_r + r_6 + R_0)}, \quad (7-52)$$

где

$$\gamma_6 = \frac{R_0}{R_r + r_6 + R_0}.$$

Чтобы найти время нарастания, нужно сначала записать изображение $K(s)$. Воспользуемся выражением (7-4а), заменив в выражении для $R_{вх}$ величину r_0 операторной величиной

$$Z_0(s) = \frac{R_0}{1 + s\tau_0},$$

где $\tau_0 = C_0 R_0$. Тогда после некоторых преобразований можно привести изображение $K_u(s)$ к виду

$$K_u(s) = K_{u0} \frac{1 + s\tau_0}{1 + \frac{s(\tau_{oe} + \gamma_5\tau_0) + s^2\gamma_5\tau_{oe}\tau_0}{1 + \beta_{oe}\gamma_6}}. \quad (7-53)$$

Анализируя оригинал этого изображения с точки зрения отсутствия выбросов [т. е. полагая $k'_u(t) = 0$ при $t = \infty$], можно показать, что такой *критический режим* соответствует условию

$$\tau_0 = C_0 R_0 = \frac{\tau_{oe}}{1 + \beta_{oe}} \quad (7-54)$$

и что время нарастания при этом выражается формулой

$$t_n \approx 2,2 \frac{\gamma_5\tau_{oe}}{1 + \beta_{oe}\gamma_6}, \quad (7-55)$$

а добротность — формулами (7-47)—(7-48), если положить в них $r_0 = 0$.

Однако нужно иметь в виду, что формулы (7-54) и (7-55), основанные на выражении (4-80а), делаются неточными при $t_n < 2,2\tau_{\alpha}$ в связи с неточностью исходного выражения (4-43) в области $t < \tau_{\alpha}$. Поэтому положим $t_n = 2,2\tau_{\alpha}$ и определим из (7-55) максимально допустимое значение γ_6 . Оно оказывается близким к $1/2$. Значит выражение (7-55) действительно при условии¹

$$R_0 \leq R_r + r_6. \quad (7-56)$$

Если выбрать C_0 больше, чем следует из (7-54), то на переходной характеристике появляются выбросы, допустимая величина которых диктуется конкретными требованиями к схеме.

¹ В случае промежуточного каскада R_r заменяется на R_K , в частности на $R_{K.онт}$ [см. (7-51), где r_0 следует положить равным нулю].

В обычных каскадах (без обратной связи) иногда используется *индуктивная коррекция фронта*, при которой катушка включается либо в коллекторную цепь последовательно с R_K (рис. 7-17, а), либо в базовую цепь параллельно входу каскада (рис. 7-17, б). Во втором варианте

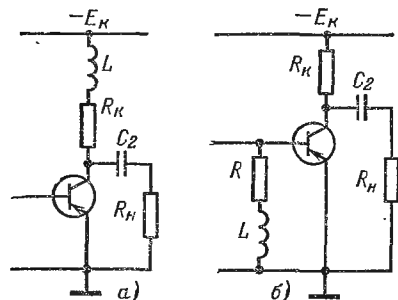


Рис. 7-17. Каскады с коллекторной (а) и базовой (б) индуктивной коррекцией.

используется не «чистая» индуктивность, а цепочка RL . Коллекторная коррекция преследует цель задержать изменение тока через R_K и тем самым форсировать нарастание тока через сопротивление нагрузки R_H . Анализ показывает [117], что таким путем удастся улучшить фронт и добротность на десятки процентов, особенно у бездрейфовых транзисторов. Базовая коррекция обеспечивает такие же начальные условия, как и без нее (поскольку в первый момент ток через катушку не идет); зато в дальнейшем корректирующая цепочка «отсасывает» часть входного тока и снижает установившееся значение выходных величин. Нетрудно заметить, что это аналогично действию емкостной коррекции (см. рис. 7-16). Однако

наличие катушки относится к существенным недостаткам такого решения, поэтому индуктивная коррекция даже в дискретных схемах используется редко; в интегральных полупроводниковых микросхемах использование индуктивности полностью исключается.

7-5. МНОГОКАСКАДНЫЕ УСИЛИТЕЛИ

При последовательном соединении нескольких каскадов главной задачей анализа является установление связи между параметрами всего усилителя и параметрами каждого из каскадов. Относительно структуры усилителей можно заметить следующее.

Иногда при построении многокаскадных усилителей разделяют два каскада ОЭ каскадом ОК, для которого, как показано в гл. 9, характерны повышенное входное и пониженное выходное сопротивления, а усиление по напряжению отсутствует. Тем самым использование «буферного» каскада ОК дает возможность искусственно увеличить сопротивление R_H для 1-го каскада ОЭ и искусственно уменьшить сопротивление R_r для 2-го каскада ОЭ. При этом коэффициент усиления такой комбинации оказывается выше, чем при непосредственном соединении двух каскадов ОЭ. Однако анализ показывает, что использование трех каскадов ОЭ (т. е. того же числа транзисторов) приводит к большему усилению, чем в случае комбинации ОЭ—ОК—ОЭ. Поэтому в дальнейшем мы анализируем только соединение однотипных каскадов.

Область средних частот. Пусть имеется N -каскадный усилитель, нагруженный на сопротивление R_H и работающий от источника сигнала с сопротивлением R_r (рис. 7-18). Коэффициент усиления такого усилителя можно записать в следующем виде:

$$K = \frac{U_{\text{вых}}}{E_r} = K_{x,x} \frac{R_H}{R_{\text{вых}} + R_H} = K_1 K_2 \dots K_{N-1} K_N \frac{R_H}{R_{\text{вых}N} + R_H}, \quad (7-57)$$

где $K_1, 2, \dots, N$ — коэффициенты усиления отдельных каскадов при холостом ходе ($K_1 = E_1/E_r$; $K_2 = E_2/E_1$; ...; $K_N = E_N/E_{N-1}$). Полагая выходные сопротивления каскадов равными R_{km} [см. (7-3)],

можно считать $R_{гm} = R_{к(m-1)}$ (за исключением 1-го каскада, для которого $R_{г1} = R_{г}$). Тогда из формулы (7-4а) при $\beta = \beta_{oe}$ и $R_{н} = \infty$ получаем для 1-го каскада:

$$K_1 = \frac{\beta_{oe1} R_{к1}}{R_{г} + R_{вх1}},$$

а для всех других каскадов

$$K_m = \frac{\beta_{oem} R_{км}}{R_{к(m-1)} + R_{вхm}}.$$

Подставляя полученные значения K в (7-57) и выделяя множитель, характеризующий источник сигнала и нагрузку, можно запи-

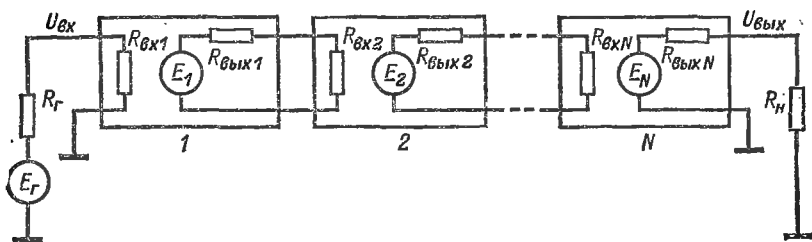


Рис. 7-18. Блок-схема многокаскадного усилителя.

сать общий коэффициент усиления в следующей компактной форме:

$$K_u = (-1)^N \frac{\beta_{oeN} (R_{н} \parallel R_{кN})}{R_{г} + R_{вх1}} \prod_{m=1}^{N-1} \beta_{oem} \frac{R_{км}}{R_{км} + R_{вх(m+1)}}. \quad (7-58a)$$

Из (7-58а) ясно, что с точки зрения усиления выгодны большие сопротивления $R_{к}$. Полагая $R_{к} \gg R_{вх}$ и считая все транзисторы одинаковыми, получаем частный случай формулы (7-58а):

$$K_u = (-1)^N \frac{\beta_{oe}^N R_{н}}{R_{г} + R_{вх}}. \quad (7-58б)$$

Область малых времен. Главная особенность многокаскадного усилителя в области малых времен состоит в том, что нагрузка каждого каскада получается комплексной, даже если внешняя нагрузка чисто активная. Это обусловлено комплексным характером входного сопротивления.

Для количественного анализа нужно подставить в (7-58а) значения $Z_{вх}(s)$, $\beta_{oe}(s)$ и $Z_{н}(s)$, из которых первые два задаются выражениями (7-35) и (7-37). Что касается нагрузки, то она в общем случае может иметь произвольный характер. Запишем ее в форме

$$Z_{н}(s) = v(s) R_{н}, \quad (7-59)$$

где $v(s)$ — дробно-рациональная функция оператора s , равная единице при $s = 0$, т. е. в установившемся режиме (на средних час-

токах). После подстановок, преобразований и сокращений получаем:

$$K_u(s) = \frac{K_{u0}v(s)}{[1 + \gamma_n(v(s) - 1)] \prod_{m=1}^N 1 + s \left(\frac{\tau_{oem}}{1 + \beta_{oem}\gamma_{6m}} \right)}, \quad (7-60)$$

где K_{u0} — коэффициент усиления на средних частотах, определяемый из выражения (7-58a), и

$$\gamma_n = \frac{R_n}{R_n + R_{kN}}.$$

Коэффициенты обратной связи γ_{6m} имеют обычный смысл [см. (7-9)], но для всех каскадов, кроме 1-го, роль R_r играет $R_{k(m-1)}$.

При активной, резистивно-емкостной и резистивно-индуктивной нагрузках оригиналами изображения (7-60) являются суммы экспоненциальных функций. В случае резистивно-емкостной нагрузки (в виде параллельного соединения R_n и C_n) функция $v(s) = 1/(1 + s\tau'_n)$, где $\tau'_n = C_n R_n$. Тогда оригиналом изображения (7-60) является переходная характеристика

$$k_u(t) = K_{u0} \sum_{m=1}^{N+1} A_m \left(1 - e^{-\frac{t}{\tau_m}} \right) \quad (7-61a)$$

(см. [62], гл. 3).

Постоянные времени в переходной характеристике (7-61a) имеют вид $\tau_m = \tau_{oem}/(1 + \beta_{oem}\gamma_{6m})$ [за исключением $\tau_{N+1} = \tau_n = C_n(R_{kN} \parallel R_n)$], а коэффициенты A_m выражаются следующим образом:

$$A_m = \frac{1}{\prod_{\substack{j=1 \\ j \neq m}}^{N+1} \left(1 - \frac{\tau_j}{\tau_m} \right)}.$$

Если входной сигнал представляет собой не ступеньку, а нарастает по экспоненциальному закону с постоянной времени $\tau_{вх}$, то в формуле (7-61a) под знаком суммы появляется еще одно слагаемое с индексом 0, причем $\tau_0 = \tau_{вх}$.

Когда все постоянные времени одинаковы, включая 1-й каскад и нагрузку, функция $k_u(t)$ упрощается:

$$k_u(t) = K_{u0} \left[1 - e^{-\frac{t}{\tau}} \sum_{m=1}^{N+1} \frac{\left(\frac{t}{\tau} \right)^{N+1-m}}{(N+1-m)!} \right]. \quad (7-61б)$$

При активной нагрузке, когда $v(s) = 1$, в выражениях (7-61) вместо $N + 1$ должно стоять N .

Выражения (7-61) имеют ту же структуру, что и в ламповых усилителях; поэтому результирующее время нарастания фронта

при резистивно-емкостной нагрузке будет определяться известным соотношением [62]

$$t_n = \sqrt{\sum_{m=1}^{N+1} t_{nm}^2} \quad (7-62a)$$

(где $t_{nm} = 2,2\tau_m$), а время задержки фронта — соотношением¹

$$t_s = \sum_{m=1}^{N+1} \tau_m. \quad (7-62b)$$

В случае одинаковых постоянных времени у всех каскадов, кроме первого (у которого положим $\tau_1 \ll \tau_m$), получается:

$$t_n = t_{nN} \sqrt{N}; \quad (7-63a)$$

$$t_s = N t_{sN}. \quad (7-63b)$$

Эти же формулы действительны тогда, когда одинаковы все каскады (включая первый), а нагрузка резистивная.

Для примера рассмотрим двухкаскадный усилитель с активной нагрузкой (рис. 7-19). Будем считать сопротивления r_K^* достаточно большими по сравнению с R_H и $R_{вх2}$. Такое допущение позволяет принять $\tau_{oe} = \tau_B$ и $\beta_{oe} = \beta$. Кроме того, положим параметры транзисторов одинаковыми, $R_T = 0$ и $R_K \gg R_{вх2}$. Тогда из (7-60), при $v(s) = 1$, получим:

$$K_u(s) = \frac{K_{u0}}{(1 + s\tau_B) \left(1 + s \frac{\tau_B}{1 + \beta\gamma_{60}}\right)}. \quad (7-64)$$

Анализ операторной части формулы (7-64) свидетельствует о том, что колебательный режим невозможен и переходный процесс носит апериодический характер. Определяющее значение имеет постоянная времени τ_B , привносимая вторым каскадом. Малая постоянная времени первого каскада $\tau_B/(1 + \beta\gamma_{60})$ проявляется в виде задержки фронта в его начальной части (см. сноску на с. 199).

В усилителях с числом каскадов больше двух и неодинаковыми постоянными времени каскадов оказываются возможными переходные процессы с выбросами. Такие выбросы называют *апериодическими*, так как они появляются вследствие наложения нескольких экспоненциальных (апериодических) функций.

Заметим, что одиночный каскад, работающий на комплексную нагрузку с емкостной реакцией, во многом походит на двухкаскад-

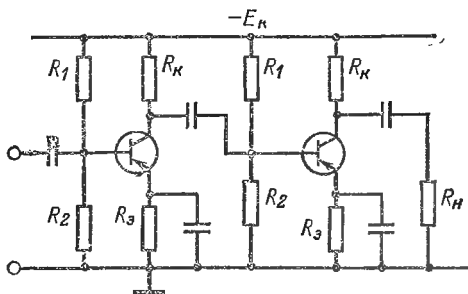


Рис. 7-19. Двухкаскадный усилитель с активной нагрузкой.

¹ При таком определении времени задержки отдельный каскад (или отдельная RC-цепочка) характеризуется значением $t_s = \tau_1$, т. е. задержка оценивается по сдвигу среднего участка фронта, а не по тому интервалу, в течение которого переходная функция близка к нулю.

ный усилитель. Разница заключается лишь в том, что и емкость, и сопротивление нагрузки могут иметь произвольное значение. Можно показать, что емкость нагрузки C_H существенна лишь тогда, когда ее значение сравнимо со значением $C_{к0}^*$, т. е. составляет десятки пикофарад и более. При меньших значениях C_H можно считать нагрузку активной, как мы и делали до сих пор.

Глава восьмая

ОБРАТНАЯ СВЯЗЬ В УСИЛИТЕЛЯХ

8-1. ВВЕДЕНИЕ

Общая теория обратной связи [118] охватывает все типы усилителей. В частности, используя матричный метод и рассматривая транзисторный усилитель как «черный ящик», характеризующийся той или иной системой параметров четырехполюсника, можно учесть наличие любого вида обратной связи [71, 119, 120]. Ценность матричного метода состоит в том, что он не связан никакими ограничениями (например, однонаправленностью сигнала), а его применение требует лишь стандартных математических операций. С другой стороны, матричному методу свойственны и серьезные недостатки: громоздкость вычисления определителей и миноров матриц в сложных схемах, а главное формальность, которая не позволяет отразить специфику транзисторов и тем самым затрудняет, например, анализ влияния различных факторов (режим, температура, частота) на свойства транзисторных усилителей. Поэтому, не исключая применения матричного метода в целом, представляется целесообразным по возможности использовать при анализе обратных связей те эквивалентные схемы и параметры, которые были приняты в предыдущих главах. Прежде чем провести такой конкретный анализ, остановимся на некоторых общих вопросах обратной связи в **о д н о н а п р а в л е н н ы х** системах¹.

8-2. ОБЩИЕ ВОПРОСЫ ОДНОНАПРАВЛЕННОЙ ОБРАТНОЙ СВЯЗИ

Основные соотношения. На рис. 8-1 в самом общем виде показана структура усилителя с обратной связью. Здесь имеются: *цепь прямой передачи*, характеризующаяся коэффициентом *прямой передачи* A (генератор величины m_2); *цепь обратной передачи*, характеризующаяся коэффициентом *обратной связи* κ (генератор величины m_4); источник сигнала G (генератор величины m_1) и нагрузка H , характеризующаяся величиной m_3 , причем $m_1 = m_{\text{вых}} = m_3$. Последнее равенство объясняется тем, что цепь обратной передачи реагирует **н е п о с р е д с т в е н н о** на величину m_3 , как правило, совпадаю-

¹ Используемые ниже термины, определения и классификация впервые обоснованы в работе [62].

щую с выходной величиной. Величины m могут быть либо напряжениями, либо токами, в зависимости от типа обратной связи (см. § 8-3).

Ответвления величин m в генераторах G , A и κ соответствуют потерям на внутренних сопротивлениях генераторов. Для идеальных генераторов (без потерь) $m_T = m_{вх}$; $m_2 = m_{вых}$ и $m_4 = m_{о.с.}$.

На рис. 8-1 взаимные направления входной величины $m_{вх}$ и величины обратной связи $m_{о.с.}$ соответствуют отрицательной обратной связи, которая будет подразумеваться и в дальнейшем.

Коэффициент прямой передачи $A = m_2/m_1$ может быть размерным или безразмерным в зависимости от величин m_2 и m_1 ; коэффициент обратной связи $\kappa = m_4/m_3$ имеет размерность, обратную размерности параметра A , или безразмерен¹.

Однонаправленность четырехполюсников A и κ отражена на рис. 8-1 тем, что зависимые генераторы действуют только на их выходах².

Пусть сначала генераторы G , A и κ идеальны (см. выше). Тогда из рис. 8-1 следует:

$$(m_T - m_{о.с.}) A = m_H.$$

Подставляя сюда $m_{о.с.} = \kappa m_H$, находим:

$$A_{о.с.} = \left(\frac{m_H}{m_T} \right)_{о.с.} = \frac{A}{1 + \kappa A}, \quad (8-1)$$

где $A_{о.с.}$ — коэффициент передачи усилителя при наличии обратной связи. Дифференцируя функцию $A_{о.с.}$ по A , нетрудно получить фундаментальное соотношение.

$$\frac{dA_{о.с.}}{A_{о.с.}} = \frac{dA/A}{1 + \kappa A}. \quad (8-2)$$

Основным свойством отрицательной обратной связи согласно (8-2) является стабилизация коэффициента A в широком

¹ Это следует из одинаковых размерностей величин m_2 и m_3 , с одной стороны, и m_1 и m_4 — с другой (последние должны иметь одинаковую размерность, поскольку они складываются друг с другом).

² В реальных случаях зависимые генераторы могут действовать как в выходной, так и во входной цепи активного четырехполюсника (см., например, рис. 4-27). В таких системах квазиоднонаправленность означает, что величина, обусловленная обратной реакцией, мала по сравнению с входной величиной либо из-за малости коэффициента обратной передачи (μ_{12} на рис. 4-27), либо из-за малости выходной величины (U_2 на рис. 4-27).

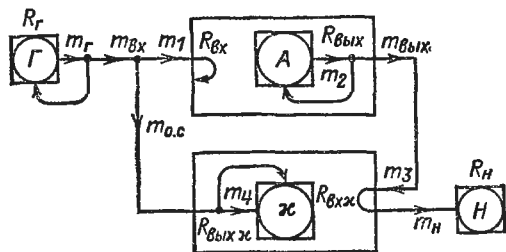


Рис. 8-1. Общая блок-схема усилителя с однонаправленной обратной связью.

смысле слова: по отношению к изменению режима активных элементов, амплитуды сигнала, частоты и т. п. Отчего бы ни изменялся коэффициент A , коэффициент $A_{o.c}$ меняется в $(1 + \kappa A)$ раз слабее.

Если глубина обратной связи $1 + \kappa A$ достаточно велика (т. е. $\kappa A \gg 1$), то коэффициент $A_{o.c}$ почти не зависит от коэффициента A :

$$A_{o.c} \approx \frac{1}{\kappa}. \quad (8-3)$$

В этом случае стабильность величины $A_{o.c}$ определяется только стабильностью коэффициента обратной связи и может быть очень высокой, когда цепь обратной передачи состоит из качественных пассивных компонентов. Если κ является комплексной или операторной величиной, то коэффициенты $A_{o.c}(\omega)$ и $A_{o.c}(s)$ согласно (8-3) будут обратными функциями по отношению к $\kappa(\omega)$ и $\kappa(s)$. Это часто используется для разного рода коррекций, а также для получения специальных частотных и переходных характеристик.

Отмеченные важные свойства отрицательной обратной связи покупаются ценой уменьшения коэффициента прямой передачи. Это ясно и из общего выражения (8-1), и из частного выражения (8-3), которое основано на неравенстве $\kappa A \gg 1$ и из которого следует: $1/\kappa \approx A_{o.c} \ll A$.

Поскольку величины A и κ в общем случае комплексные (или операторные), то с ростом частоты сигнала (или скорости его изменения) неравенство $\kappa A \gg 1$ рано или поздно нарушается и коэффициент $A_{o.c}$ начинает зависеть от A . При достаточно высокой частоте (или в области очень малых времен) в с е г д а получается обратное неравенство $\kappa A \ll 1$ и тогда $A_{o.c} \approx A$, т. е. стабилизирующее действие отрицательной обратной связи исчезает (в частности, сливаются частотные и переходные характеристики коэффициентов $A_{o.c}$ и A).

В р е а л ь н ы х генераторах (с конечными внутренними сопротивлениями) влияние входных сопротивлений четырехполюсников A и κ , а также сопротивления нагрузки можно учесть, не конкретизируя схемы на рис. 8-1. А именно положим, что в реальном случае на вход четырехполюсника A попадает доля ξ_{sx} величины m_r и доля ξ_3 величины m_4 , а в нагрузку (и тем самым на вход четырехполюсника κ) — доля $\xi_{вых}$ величины m_2 [коэффициенты входа и выхода $\xi_{вх}$ и $\xi_{вых}$ уже использовались в формулах (7-4) и (7-7)]. Поскольку входная цепь — общая для генераторов Γ и κ и поскольку величины m_r и m_4 передаются на один и тот же участок этой цепи, имеет место равенство $\xi_4 = \xi_{вых}$.

В результате можно записать соотношения:

$$m_1 = m_r \xi_{вх} - m_4 \xi_{вх};$$

$$m_4 = \kappa m_3 = \kappa m_2 \xi_{вых} = \kappa A m_1 \xi_{вых};$$

$$m_4 = m_{вых} = m_2 \xi_{вых} = A m_1 \xi_{вых},$$

из которых первые два позволяют найти m_1 , а третье после подстановки m_1 дает выражение

$$A'_{o.c} = \left(\frac{m_n}{m_r} \right)_{o.c} = \frac{A'}{1 + \kappa A'}, \quad (8-4a)$$

совпадающее по форме с (8-1), но с коэффициентом прямой передачи

$$A' = \frac{m_n}{m_r} = A \xi_{вх} \xi_{вых}. \quad (8-4б)$$

Таким образом, общие выражения (8-1)—(8-3) действительны для любой реальной однонаправленной системы, если характеризовать усилительную способность отношением величины m_n в нагрузке к величине m_r в генераторе сигнала (что было обосновано в сноске на с. 330).

Помимо передаточного параметра всегда представляют интерес входное и выходное сопротивления усилителя с обратной связью. Зависимость этих сопротивлений от глубины обратной связи также можно исследовать в общем виде, не конкретизируя схемы на рис. 8-1.

Напомним, что входное сопротивление определяется величиной $m_{вх}$ (а не m_r), т. е. генератор Γ при оценке $R_{вх}$ нужно считать идеальным¹. Тогда

$$\begin{aligned} m_1 &= m_{вх} \xi'_{вх} - m_4 \xi'_{вх}; \\ m_4 &= \kappa A m_1 \xi_{вых}, \end{aligned}$$

где $\xi'_{вх} > \xi_{вх}$, поскольку генератор Γ считается идеальным. Из приведенных соотношений следует:

$$(m_1)_{o.c} = \frac{m_{вх} \xi_{вх}^*}{1 + \kappa A \xi_{вх} \xi_{вых}}. \quad (8-5)$$

Поскольку в отсутствие цепи обратной передачи $m_1 = m_{вх}$, приходим к выводу, что наличие цепи κ приводит к уменьшению величины m_1 по двум причинам: из-за конечного выходного сопротивления генератора κ (т. е. в связи с тем, что $\xi'_{вх} < 1$), а главное, благодаря обратной связи, характеризуемой глубиной $1 + \kappa A \xi_{вх} \xi_{вых}$. Таким образом, если величины $m_{вх}$ и m_1 — напряжения, то благодаря обратной связи уменьшается напряжение U_1 , что приводит к уменьшению тока I_1 и соответственно увеличению эквивалентного входного сопротивления (определяемого как $U_{вх}/I_1$). Если же величины $m_{вх}$ и m_1 — токи, то благодаря обратной связи уменьшается ток I_1 , что приводит к уменьшению напряжения U_1 и соответственно увеличению эквивалентной входной проводимости ($I_{вх}/U_1$), а эквивалентное входное сопротивление ($U_1/I_{вх}$) уменьшается.

Для того чтобы оценить выходное сопротивление, следует учесть, что оно определяется величиной $m_{вых}$, задаваемой в отсутствие нагрузки и входного сигнала ($m_r = 0$). Влияние выходного сопротивления проявляется в различии величин m_2 и $m_{вых}$, поэтому рассмотрим зависимость разности $m_2 - m_{вых}$ от глубины обратной связи. Из рис. 8-1 видно, что в отсутствие R_n его роль играет входное сопротивление четырехполосника κ . Поэтому, чтобы оценить выходное сопротивление усилителя «в чистом виде», опустим не только R_n , но и R_κ . Тогда, задав извне величину $m_{вых}$, можем записать:

$$m_2 = -\kappa \xi_{вх} A m_{вых}$$

и, следовательно,

$$(m_2 - m_{вых})_{o.c} = -m_{вых} (1 + \kappa A \xi_{вх}). \quad (8-6)$$

¹ См. предпоследний абзац в разделе «Внутренняя обратная связь», § 7-2.

Поскольку в отсутствие цепи обратной передачи $m_2 = 0$, т. е. $m_2 = m_{\text{вых}} = -m_{\text{вых}}$, приходим к выводу, что наличие цепи κ приводит к увеличению разности $m_2 - m_{\text{вых}}$. Таким образом, если величины m_2 и $m_{\text{вых}}$ — напряжения, то увеличение разности $U_2 - U_{\text{вых}}$ (благодаря обратной связи) влечет увеличение тока через внутреннее сопротивление генератора U_2 , а вместе с тем и выходного тока. При заданной э. д. с. $U_{\text{вых}}$ это равносильно уменьшению эквивалентного выходного сопротивления (определяемого как $-U_{\text{вых}}/I_{\text{вых}}$). Если же величины m_2 и $m_{\text{вых}}$ — токи, то увеличение разности $I_2 - I_{\text{вых}}$ (благодаря обратной связи) влечет увеличение напряжения на

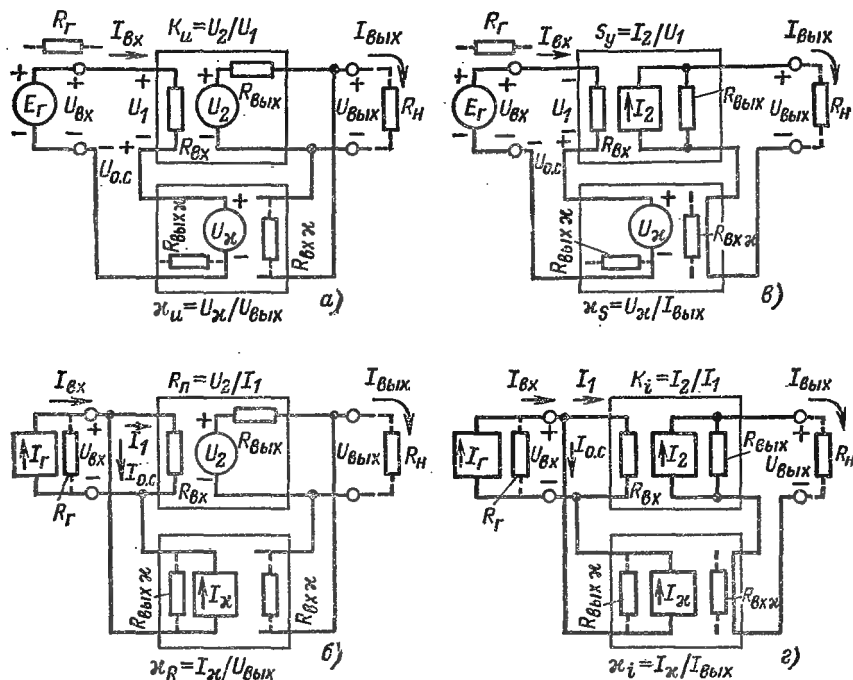


Рис. 8-2. Скелетные схемы усилителей с обратной связью.

а — последовательной по напряжению; б — параллельной по напряжению; в — последовательной по току; г — параллельной по току.

внутреннем сопротивлении генератора I_2 , а вместе с тем и на выходе. При заданном токе $I_{\text{вых}}$ это равносильно уменьшению эквивалентной выходной проводимости (определяемой как $-I_{\text{вых}}/U_{\text{вых}}$) и увеличению эквивалентного выходного сопротивления.

В обоих случаях в величину эквивалентного выходного сопротивления можно внести поправку на входное сопротивление цепи κ , которым раньше пренебрегли. Поправка состоит в параллельном или последовательном соединении обоих сопротивлений, в зависимости от типа обратной связи — по напряжению или по току, см. рис. 8-2. Результирующее сопротивление оказывается таким же, как если бы мы с самого начала объединили выходное сопротивление генератора A с входным сопротивлением цепи κ , а влияние обратной связи отразили соответственно делением или умножением суммарного сопротивления на глубину обратной связи $1 + \kappa A'_{\text{вх}} \xi'_{\text{вых}}$. Здесь $\xi'_{\text{вых}}$ — коэффициент выхода, аналогичный коэффициенту $\xi_{\text{вых}}$, но отличающийся от него, поскольку нагрузка отключена и на выходе имеется лишь входное сопротивление цепи обратной передачи; поэтому $\xi'_{\text{вых}} > \xi_{\text{вых}}$.

Классификация. На рис. 8-2 представлено четыре основных способа реализации обратной связи. Все они соответствуют общей структурной схеме на рис. 8-1.

Обратная связь по напряжению (рис. 8-2, а и б) характерна тем, что сигнал обратной связи ($U_{o.c}$ или $I_{o.c}$) пропорционален выходному напряжению.

Обратная связь по току (рис. 8-2, в и г) характерна тем, что сигнал обратной связи ($U_{o.c}$ или $I_{o.c}$) пропорционален выходному току.

По способу сложения сигнала обратной связи с входным сигналом различают: обратную связь со сложением напряжений (рис. 8-2, а и в) и обратную связь со сложением токов (рис. 8-2, б и г). В литературе обратную связь со сложением напряжений часто называют последовательной, а со сложением токов — параллельной. Эти названия менее точны по физическому смыслу (и не всегда правильно используются), но удобны в силу их кратности¹.

Исходя из размерностей выходных и входных величин, приходим к выводу, что коэффициентами передачи усилителей, показанных на рис. 8-2, должны быть: коэффициент усиления напряжения K_u (рис. 8-2, а), сопротивление передачи R_n (рис. 8-2, б), крутизна усилителя S_y (рис. 8-2, в) и коэффициент усиления тока K_i (рис. 8-2, г)². При наличии обратной связи все параметры подчиняются общим закономерностям (8-1)—(8-4), причем коэффициенты обратной связи κ_n и κ_i безразмерны, а коэффициенты κ_R и κ_S имеют соответственно размерности проводимости и сопротивления.

Коэффициенты ξ (см. предыдущий раздел), характеризующие распределение напряжений и токов на входе и выходе усилителя при бездействующей обратной связи ($\kappa = 0$), зависят от структуры входа и выхода схемы. А именно, при сложении напряжений на входе (рис. 8-2, а и в)

$$\xi_{вх u} = \frac{U_1}{E_r} = \frac{R_{вх}}{R_{вх} + R_r + R_{вх \kappa}}; \quad (8-7a)$$

$$\xi'_{вх u} = (\xi_{вх u})_{R_r = 0} = \frac{R_{вх}}{R_{вх} + R_{вх \kappa}}; \quad (8-7б)$$

при сложении токов на входе (рис. 8-2, б и г)

$$\xi_{вх i} = \frac{I_1}{I_r} = \frac{g_{вх}}{g_{вх} + g_r + g_{вх \kappa}}; \quad (8-7в)$$

$$\xi'_{вх i} = (\xi_{вх i})_{g_r = 0} = \frac{g_{вх}}{g_{вх} + g_{вх \kappa}}. \quad (8-7г)$$

¹ Иногда [71, 121, 122] используют аналогичную терминологию и для выходной цепи. Тогда схемы на рис. 8-2, а—г, называют соответственно схемами с последовательно-параллельной, параллельной, последовательной и параллельно-последовательной обратной связью.

² Крутизну усилителя было бы логичнее назвать «проводимостью передачи», но этот термин не имеет распространения в усилительной технике.

При обратной связи по напряжению (рис. 8-2, а и б)

$$\xi_{\text{вых} u} = \frac{U_n}{U_2} = \frac{g_{\text{вых}}}{g_{\text{вых}} + g_n + g_{\text{вх} \kappa}}; \quad (8-8a)$$

$$\xi'_{\text{вых} u} = (\xi_{\text{вых} u})_{g_n = 0} = \frac{g_{\text{вых}}}{g_{\text{вых}} + g_{\text{вх} \kappa}}; \quad (8-8б)$$

при обратной связи по току (рис. 8-2, в и г)

$$\xi_{\text{вых} i} = \frac{I_n}{I_2} = \frac{R_{\text{вых}}}{R_{\text{вых}} + R_n + R_{\text{вх} \kappa}}; \quad (8-8в)$$

$$\xi'_{\text{вых} i} = (\xi_{\text{вых} i})_{R_n = 0} = \frac{R_{\text{вых}}}{R_{\text{вых}} + R_{\text{вх} \kappa}}. \quad (8-8г)$$

Рассмотрим входные и выходные сопротивления усилителя с обратной связью. Естественно, что входное сопротивление зависит от структуры входной цепи, а выходное сопротивление — от структуры выходной цепи. Исходя из соотношений (8-5), (8-6) и рис. 8-2, можно получить следующие выражения:

при сложении напряжений на входе (рис. 8-2, а и в)

$$R_{\text{вх. о. с}} = (R_{\text{вх}} + R_{\text{вх} \kappa}) (1 + \kappa A \xi'_{\text{вх} u} \xi_{\text{вых} u}); \quad (8-9a)$$

при сложении токов на входе (рис. 8-2, б и г)

$$g_{\text{вх. о. с}} = (g_{\text{вх}} + g_{\text{вх} \kappa}) (1 + \kappa A \xi'_{\text{вх} i} \xi_{\text{вых} i}); \quad (8-9б)$$

при обратной связи по напряжению (рис. 8-2, а и б)

$$g_{\text{вых. о. с}} = (g_{\text{вых}} + g_{\text{вх} \kappa}) (1 + \kappa A \xi_{\text{вх} u} \xi'_{\text{вых} u}); \quad (8-10a)$$

при обратной связи по току (рис. 8-2, в и г)

$$R_{\text{вых. о. с}} = (R_{\text{вых}} + R_{\text{вх} \kappa}) (1 + \kappa A \xi_{\text{вх} i} \xi'_{\text{вых} i}), \quad (8-10б)$$

Как видим, при последовательной обратной связи входное сопротивление увеличивается, а при параллельной — уменьшается. Что касается выходного сопротивления, то при обратной связи по напряжению выходное сопротивление уменьшается, а при обратной связи по току — увеличивается.

Меньшая величина $R_{\text{вых}}$, естественно, означает меньшее влияние нагрузки на выходное напряжение. Поэтому говорят, что обратная связь по напряжению стабилизирует выходное напряжение. По аналогичным причинам обратная связь по току стабилизирует выходной ток.

Из выражений (8-4), (8-9) и (8-10) следует, что при малых коэффициентах входа и выхода труднее обеспечить глубокую обратную связь, а значит, высокую стабильность усиления и существенное преобразование входных и выходных сопротивлений. Кроме того, если малые значения ξ обусловлены большими значениями R_{Γ} , R_n или g_{Γ} , g_n [см. (8-7), (8-8)], то, как легко убедиться, параметры усилителя будут в той или иной мере зависеть от внешних сопротивлений R_{Γ} и R_n , что обычно нежелательно. Поэтому каждому из четырех основных типов усилителей [62] — усилителям напря-

жения ($R_{\Gamma} \ll R_{\text{вх}}$), усилителям тока ($R_{\Gamma} \gg R_{\text{вх}}$), усилителям с потенциальным выходом ($R_{\text{н}} \gg R_{\text{вых}}$) и усилителям с токовым выходом ($R_{\text{н}} \ll R_{\text{вых}}$) — соответствует свой оптимальный тип обратной связи, для которого коэффициенты $\xi_{\text{вх}}$ и $\xi_{\text{вых}}$ близки к единице.

А именно, учитывая структуру формул (8-7) и (8-8), последовательная обратная связь свойственна усилителям напряжения, параллельная обратная связь — усилителям тока, обратная связь по напряжению — усилителям с потенциальным выходом и обратная связь по току — усилителям с токовым выходом. Только при таком соответствии отрицательная обратная связь оказывается наиболее эффективной, т. е. обеспечивает в полной мере свое стабилизирующее действие.

Если же, например, в усилителе с токовым выходом использовать обратную связь по напряжению, то выходное сопротивление согласно (8-10а) уменьшится и изменения нагрузки $R_{\text{н}}$ будут сильнее влиять на значение выходного тока, что не отвечает назначению такого усилителя. В пределе, при $R_{\text{н}} \approx 0$, обратная связь по напряжению в усилителях с токовым выходом делается вообще невозможной. Если в усилителе напряжения осуществить параллельную обратную связь, то входное сопротивление согласно (8-9б) уменьшится и усилитель будет сильнее нагружать источник сигнала, что весьма нежелательно в усилителях напряжения. В пределе, при $R_{\Gamma} \approx 0$, параллельная обратная связь вообще бессмысленна, так как при этом ток обратной связи полностью замыкается через источник сигнала и не попадает на вход усилителя. Соответственно коэффициент усиления будет примерно такой же, как и без обратной связи.

В табл. 8-1 приведены основные выражения, свойственные каждой из схем на рис. 8-2. В тех случаях, когда коэффициент прямой передачи не соответствует структуре усилителя (например, когда для схемы на рис. 8-2, б используется не сопротивление передачи, а коэффициент усиления тока или напряжения), можно воспользоваться формулами пересчета, приведенными в табл. 8-2.

Эти формулы получены путем элементарных преобразований, основанных на соотношениях $E_{\Gamma} = I_{\Gamma} R_{\Gamma}$ (принцип взаимного соответствия или дуальности [62]) и $U_{\text{вых}} = I_{\text{вых}} R_{\text{н}}$.

Полученные выражения (табл. 8-1) имеют од н о т и п н у ю простую структуру и действительны для всех случаев, когда цепи обратной передачи однонаправленны или могут быть приведены к однонаправленным без принципиальных погрешностей¹.

Если же однонаправленность цепи обратной передачи выражена слабо, то выражения (8-1) — (8-10) нуждаются в существенных поправках и теряют свою простую и физически наглядную структуру. Именно в этом случае приходится прибегать к матричному методу (см. § 8-1). В последующих параграфах рассматриваются примеры обратных связей, анализ которых может основываться на изложенной выше основе, т. е. квазиоднонаправленных.

¹ Отсутствие однонаправленности и цепи прямой передачи означает наличие в ней внутренней обратной связи. В этом случае можно разбить четырехполюсник A на два однонаправленных или квазиоднонаправленных (A' и A''), см. § 8-3.

Основные соотношения при разных типах обратной связи¹

Вход	Выход	
	Обратная связь по напряжению	Обратная связь по току
Обратная связь со сложением напряжений (последовательная)	$K_{ио.с} = \frac{K_{и}}{1 + \kappa_{и} K_{и}};$ $R_{вх.о.с} = (R_{вх} + R_{выхи}) \times (1 + \kappa_{и} A_{и} \xi'_{вхи} \xi_{выхи});$ $g_{вых.о.с} = (g_{вых} + g_{выхи}) \times (1 + \kappa_{и} A_{и} \xi_{вхи} \xi'_{выхи});$	$S_{у.о.с} = \frac{S_{у}}{1 + \kappa_{с} S_{у}};$ $R_{вх.о.с} = (R_{вх} + R_{выхи}) \times (1 + \kappa_{с} A_{с} \xi'_{вхи} \xi_{выхи});$ $R_{вых.о.с} = (R_{вых} + R_{вхи}) \times (1 + \kappa_{с} A_{с} \xi_{вхи} \xi'_{выхи});$
Обратная связь со сложением токов (параллельная)	$R_{п.о.с} = \frac{R_{п}}{1 + \kappa_{R} R_{п}};$ $g_{вх.о.с} = (g_{вх} + g_{выхи}) \times (1 + \kappa_{R} A_{п} \xi'_{вхи} \xi_{выхи});$ $g_{вых.о.с} = (g_{вых} + g_{вхи}) \times (1 + \kappa_{R} A_{п} \xi_{вхи} \xi'_{выхи});$	$K_{ио.с} = \frac{K_{и}}{1 + \kappa_{i} K_{и}};$ $g_{вх.о.с} = (g_{вх} + g_{выхи}) \times (1 + \kappa_{i} A_{i} \xi'_{вхи} \xi_{выхи});$ $R_{вых.о.с} = (R_{вых} + R_{вхи}) \times (1 + \kappa_{i} A_{i} \xi_{вхи} \xi'_{выхи});$

¹ В табл. 8-1 параметры $K_{и}$, $R_{п}$, $S_{у}$ и K_{i} связаны с коэффициентами цепей прямой передачи соотношениями (8-46):

$$K_{и} = A_{и} \xi_{вхи} \xi_{выхи}; \quad S_{у} = A_{с} \xi_{выхи} \xi_{вхи}; \quad R_{п} = A_{R} \xi_{вхи} \xi_{выхи}; \quad K_{i} = A_{i} \xi_{вхи} \xi_{выхи}.$$

Таблица 8-2

Соотношения между передаточными параметрами

Передаточный параметр	Выражение данного передаточного параметра через			
	$K_{и}$	$S_{у}$	$R_{п}$	K_{i}
$K_{и}$	—	$S_{у} R_{п}$	$R_{п}/R_{г}$	$K_{i} (R_{и}/R_{г})$
$S_{у}$	$K_{и}/R_{и}$	—	$R_{п}/(R_{г} R_{и})$	$K_{i}/R_{г}$
$R_{п}$	$K_{и}/R_{г}$	$S_{у} (R_{г} R_{и})$	—	$K_{i}/R_{и}$
K_{i}	$K_{и} (R_{г}/R_{и})$	$S_{у} R_{г}$	$R_{п}/R_{и}$	—

8-3. ВНУТРЕННЯЯ ОБРАТНАЯ СВЯЗЬ

С внутренней обратной связью мы уже познакомились в § 7-2 и отразили ее во всех основных формулах, описывающих RC-усилители. В данном параграфе мы установим тип этой обратной связи в соответствии с общими определениями¹.

¹ Ниже не учитывается обратная связь по напряжению, т. е. коэффициент $\mu_{э.к.}$. Об этом виде внутренней связи будет сказано в конце параграфа.

Источником сигнала — генератор тока. Если управляющей величиной в каскаде ОЭ считать ток базы и если учесть, что в выходной цепи эквивалентной схемы действует тоже генератор тока, то внутреннюю обратную связь в таком каскаде логично рассматривать как обратную связь по току со сложением токов на входе (рис. 8-2, *г*). Для простоты положим суммарную нагрузку $R_k \parallel R_n$ равной нулю и представим источник сигнала генератором тока I_r . Тогда эквивалентная схема каскада ОЭ изобразится так, как показано на рис. 8-3, *а*. Она имеет структуру, показанную на рис. 8-2, *г*, с тем небольшим отличием, что последовательно с источником сигнала оказывается включенным сопротивление r_b .

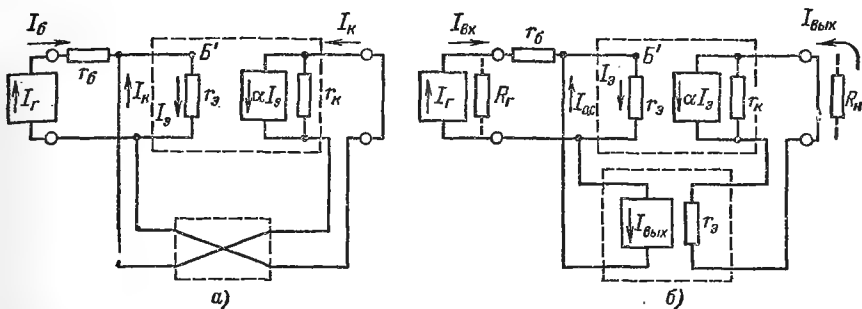


Рис. 8-3. Эквивалентная схема ОЭ.

а — приведенная к структуре на рис. 8-2, *г*; *б* — с явно выраженным четырехполюсником обратной связи.

Как видим, преобразование эквивалентной схемы ОЭ к «канонической» структуре, соответствующей параллельной связи по току, привело к необходимости использовать параметры, свойственные включению ОБ¹. Легко заметить, что обратная связь в данном случае положительна, т. е. способствует увеличению усилительного параметра цепи прямой передачи (в данном случае — коэффициента α), но одновременно снижает его стабильность. Последний вывод подтверждается тем, что коэффициент α при включении ОБ более стабилен, чем коэффициент β при включении ОЭ [см. замечания к формуле (4-68) на с. 214].

Коэффициентом прямой передачи в схеме рассматриваемого типа является коэффициент усиления тока K_I . Считая четырехполюсник обратной связи однонаправленным, представим схему так, как показано на рис. 8-3, *б*. Следовательно, в данном случае $R_{вх} = r_b$; $g_{вых} = 0$. Из рис. 8-3, *б* получаем параметры цепи прямой передачи:

$$A_I = -\alpha; \quad (8-11a)$$

$$g_{вх} = g_s = 1/r_b; \quad (8-11б)$$

$$R_{вых} = r_k. \quad (8-11в)$$

¹ Действительно, после преобразования схемы входным сопротивлением цепи прямой передачи оказалось сопротивление r_b , а следовательно, управляющим током в этой цепи должен быть ток эмиттера.

В целом, с учетом обратной связи ($\alpha = 1$) и при условии $R_{\Gamma} = \infty$ и $R_{\Pi} = 0$, нужно (8-11а) разделить, а (8-11б) умножить на глубину обратной связи (см. табл. 8-1), в данном случае на величину $1 - \alpha$. Тогда параметры транзистора принимают знакомый вид, свойственный включению ОЭ:

$$K_I \approx \frac{-\alpha}{1-\alpha} = -\beta; \quad (8-12a)$$

$$g_{\text{вх}} = (1 - \alpha) g_{\text{э}} = \frac{1 - \alpha}{r_{\text{э}}}; \quad (8-12б)$$

$$R_{\text{вых}} \approx r_{\text{к}} (1 - \alpha) = r_{\text{к}}^*. \quad (8-12в)$$

К величине $R_{\text{вх}} = r_{\text{э}} (1 + \beta)$ следует прибавить сопротивление $r_{\text{б}}$, поскольку истинный вход транзистора на рис. 8-3, б расположен левее зажимов обратной связи.

Влияние сопротивлений R_{Γ} и R_{Π} , показанных пунктиром, учитывается с помощью коэффициентов $\xi_{\text{вх}i}$ и $\xi_{\text{вых}i}$, а влияние сопротивления $r_{\text{б}}$ — путем преобразования компонентов I_{Γ} , R_{Γ} , $r_{\text{б}}$ по теореме об эквивалентном генераторе. Опуская выкладки (см. 3-е издание данной книги), запишем получающиеся выражения для параметров каскада:

$$K_I = - \frac{\beta_{\text{oe}}}{1 + \beta_{\text{oe}} \gamma_{\text{б}}} \frac{R_{\Gamma}}{R_{\Gamma} + r_{\text{б}} + r_{\text{э}}}; \quad (8-13a)$$

$$R_{\text{вх}} = (r_{\text{б}} + r_{\text{э}}) (1 + \beta_{\text{oe}} \gamma_{\text{б}}); \quad (8-13б)$$

$$R_{\text{вых}} = r_{\text{к}}^* (1 + \beta \gamma_{\text{б}}). \quad (8-13в)$$

Величины β_{oe} , $\gamma_{\text{б}}$, $\gamma_{\text{б}}$ имеют тот же смысл, что и в § 7-2. Если, как обычно бывает, результирующая нагрузка состоит из параллельно включенных сопротивлений $R_{\text{к}}$ и R_{Π} , то (8-13а) следует умножить на $R_{\text{к}} / (R_{\text{к}} + R_{\Pi})$, а (8-13в) сложить параллельно с $R_{\text{к}}$. После этого выражения (8-13) совпадут с (7-17), (7-15) и (7-22).

Структура схемы (рис. 8-3) сохраняется и для каскада ОБ, поскольку эквивалентные схемы транзистора в обоих включениях структурно одинаковы (ср. рис. 4-13 и 4-24). Поэтому для каскада ОБ достаточно на рис. 8-3 и в предыдущих формулах заменить $r_{\text{э}}$ на $r_{\text{б}}$, $r_{\text{к}}$ на $r_{\text{к}}^*$ и $\alpha I_{\text{б}}$ на $-\beta I_{\text{б}}$. Тогда вместо формул (8-12) получим для каскада ОБ:

$$K_I = \frac{\beta}{1 + \beta} = \alpha; \quad g_{\text{вх}} = (1 + \beta) g_{\text{б}}; \quad R_{\text{вых}} = r_{\text{к}}^* (1 + \beta) = r_{\text{к}}$$

(к величине $R_{\text{вх}} = r_{\text{б}} / (1 + \beta)$ нужно прибавить сопротивление $r_{\text{э}}$, включенное до четырехполюсника прямой передачи).

Обратная связь в данном случае отрицательная, а цепь прямой передачи представляет собой эквивалентную схему транзистора при включении ОЭ.

Источник сигнала — генератор э. д. с. На рис. 8-4, а показана идеализированная эквивалентная схема каскада ОЭ, структура которой соответствует последовательной обратной связи по току (см. рис. 8-2, в). Источником сигнала является генератор

E_r , а коэффициентом передачи — крутизна усилителя. Обратная связь в данном случае отрицательная, т.е. она уменьшает коэффициент передачи, но повышает стабильность усиления.

Считая четырехполюсник обратной связи однонаправленным, представим его так, как показано на рис. 8-4, б. В данном случае

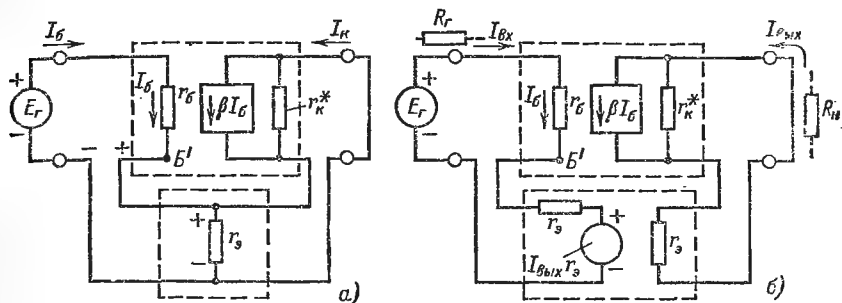


Рис. 8-4. Эквивалентная схема ОЭ.

а — приведенная к структуре на рис. 8-2, а; б — с явно выраженным четырехполюсником обратной связи.

$R_{вх\kappa} = R_{вых\kappa} = r_3$. Из рис. 8-4, б следует, что параметры цепи прямой передачи равны:

$$A_S = -\frac{\beta I_6}{U_{r_6}} = -\frac{\beta}{r_6}; \quad (8-14a)$$

$$R_{вх} = r_6; \quad (8-14б)$$

$$R_{вых} = r_k^*. \quad (8-14в)$$

Поскольку сопротивление r_3 внутреннее, т.е. «неизбежное», то даже при бездействующей обратной связи величину (8-14а) нужно умножить на «поправочные» коэффициенты $\xi'_{вх\kappa}$ и $\xi'_{вых\kappa}$, а величины (8-14б) и (8-14в) сложить соответственно с $R_{вх\kappa}$ и $R_{вых\kappa}$ (см. табл. 8-1). Используя (8-7б) и (8-8г), получаем:

$$\xi'_{вх\kappa} = \frac{r_6}{r_6 + r_3}; \quad \xi'_{вых\kappa} = \frac{r_k^*}{r_k^* + r_3} \approx 1.$$

Следовательно, параметры транзистора без учета внутренней обратной связи имеют вид:

$$S = -\frac{\beta}{r_6} \xi'_{вх\kappa} \xi'_{вых\kappa} \approx -\frac{\beta}{r_6 + r_3}; \quad (8-15a)$$

$$R_{вх} = r_6 + r_3; \quad (8-15б)$$

$$R_{вых} = r_k^* + r_3. \quad (8-15в)$$

С учетом же внутренней обратной связи ($\kappa_S = -r_3$) нужно (8-15а) разделить, а (8-15б) и (8-15в) умножить на глубину обратной связи (см. табл. 8-1), в данном случае на величину $1 + \beta \gamma_{60}$. Поэтому параметры реального транзистора при-

нимают вид ¹:

$$S = - \frac{\beta}{(r_6 + r_3)(1 + \beta \gamma_{60})} = - \frac{\beta}{r_6 + r_3(1 + \beta)}; \quad (8-16a)$$

$$R_{вх} = (r_6 + r_3)(1 + \beta \gamma_{60}); \quad (8-16б)$$

$$R_{вых} = r_K^* (1 + \beta \gamma_{60}). \quad (8-16в)$$

Последние два выражения знакомы по § 7-2.

Чтобы учесть влияние сопротивлений R_r и R_n , показанных на рис. 8-4, б пунктиром, нужно воспользоваться коэффициентами $\xi_{вх II}$ и $\xi_{вых I}$, которые согласно (8-7а) и (8-8а) имеют значения:

$$\xi_{вх II} = \frac{r_6}{r_6 + R_r + r_3}; \quad \xi_{вых I} = \frac{r_K^*}{r_K^* + R_n + r_3}.$$

Тогда с помощью табл. 8-1 и формул (8-14) получим параметры каскада ОЭ:

$$S_y = - \frac{\beta_{oe}}{(R_r + r_6 + r_3)(1 + \beta_{oe} \gamma_6)}; \quad (8-17a)$$

$$R_{вх} = (r_6 + r_3)(1 + \beta_{oe} \gamma_6); \quad (8-17б)$$

$$R_{вых} = r_K^* (1 + \beta \gamma_6), \quad (8-17в)$$

где γ_6 имеет значение (7-9), а β_{oe} отличается от значения (7-20) лишь тем, что вместо суммарного сопротивления $R_K \parallel R_n$ учтено одно сопротивление R_n .

Сравнительная оценка. Таким образом, при использовании Т-образной эквивалентной схемы *внутренняя обратная связь в каскаде ОЭ всегда является обратной связью по току*. Что касается классификации по входу, то в зависимости от типа источника сигнала (генератор тока или генератор э. д. с.) внутренняя обратная связь может быть либо параллельной (сложение токов, рис. 8-3), либо последовательной (сложение напряжений, рис. 8-4). В первом случае основным усилительным параметром будет коэффициент усиления тока, во втором — крутизна.

Первый случай (источник сигнала в виде генератора тока) характерен для многих транзисторных каскадов. Однако он связан с использованием параметров α и r_K , неудобных для схемы с общим эмиттером; кроме того, приходится иметь дело с положительной обратной связью, что менее привычно в усилительной технике; наконец в инженерной практике более распространены расчеты с коэффициентом усиления напряжения, чем с коэффициентом усиления тока. Поэтому при анализе внутренней обратной связи методически удобнее исходить из эквивалентной схемы с последовательной связью и генератором э. д. с. в качестве источника сигнала (рис. 8-4).

В тех случаях, когда сопротивление R_r настолько велико, что пользование «естественным» параметром такой схемы — крутиз-

¹ Для идеального транзистора, у которого $r_6 = 0$, получается $S = -\alpha/r_3$, что совпадает с определением крутизны на с. 222 с точностью до знака.

ной нелогично, можно перейти к коэффициенту усиления тока (см. табл. 8-2):

$$K_I = S_y R_r = - \frac{\beta_{oe}}{1 + \beta_{oe} \gamma_6} \frac{R_r}{R_r + r_6 + r_s}.$$

Последнее выражение, естественно, совпадает с (8-13а). При условии $R_r \gg r_6 + r_s$ оба они дают значение

$$K_I \approx - \frac{\beta_{oe}}{1 + \beta_{oe} \gamma_6}, \quad (8-18)$$

которое является основой важнейших формул в гл. 6 и 7. Эту величину можно трактовать как коэффициент усиления тока транзисторного каскада при достаточно большом сопротивлении источника сигнала и достаточно малом сопротивлении нагрузки.

Анализ и выводы, приведенные в данном параграфе, относятся, конечно, не только к области средних частот, но и к другим частотным или временным диапазонам, а также к постоянным составляющим токов и напряжений. Именно такой была методика исследований в гл. 6 и 7.

В заключение напомним, что наряду с внутренней обратной связью по току в транзисторных каскадах всегда действует внутренняя обратная связь по напряжению, которая на рис. 4-13 и 4-24 отражена генератором э. д. с. $\mu_{эк} U_K$. Таким образом, внутренняя обратная связь, вообще говоря, смешанная. Однако роль обратной связи по напряжению, как правило, невелика (см. § 7-1) и ее приходится учитывать лишь в каскадах с весьма высокоомной нагрузкой. Методика учета была показана в § 7-2 [см. формулы (7-21)].

8-4. ОБРАТНАЯ СВЯЗЬ ПО ПЕРЕМЕННОМУ ТОКУ

Один из важнейших каскадов с глубокой отрицательной обратной связью — эмиттерный повторитель — подробно рассматривается в гл. 9. В настоящем параграфе рассмотрены другие варианты обратных связей, как местных, так и общих. При этом во всех случаях обратная связь считается глубокой.

Местная обратная связь по току. Этот вид обратной связи, знакомый по каскаду с коррекцией (см. рис. 7-16), обеспечивается простейшим способом — включением активного сопротивления R_0 последовательно с эмиттером (рис. 8-5). При этом остаются в силе выражения (8-17), а также выражения, приведенные в § 7-2 и 7-4, с заменой величины r_s на сумму $r_s + R_0$. На практике всегда берут $R_0 \gg r_s$, чтобы полнее использовать преимущества обратной связи¹, поэтому в указанные выражения можно вместо r_s подставлять R_0 .

¹ Казалось бы, глубину обратной связи можно увеличить, не включая внешний резистор R_0 , а уменьшая рабочий ток транзистора, поскольку уменьшение тока I_s приводит согласно (4-22) к увеличению сопротивления r_s . Однако такой вариант неравноценен основному (рис. 8-5): во-первых, глубина обратной связи оказывается функцией режима (рабочего тока); во-вторых, режим малых токов (микрорежим) характерен ухудшением усилительных и частотных свойств схемы, а также меньшим динамическим диапазоном.

Произведя эту замену в формуле (8-17б), получим:

$$R_{вх} = (r_6 + R_0) (1 + \beta_{oe} \gamma'_{60}) \approx \beta_{oe} R_0, \quad (8-19)$$

где

$$\gamma'_{60} \approx \frac{R_0}{r_6 + R_0};$$

$$\beta_{oe} = \beta \gamma'_k = \beta \frac{r_k^*}{r_k^* + R_k \parallel R_n}$$

[коэффициент γ'_k записан с учетом условия $R_k \parallel R_n \gg R_0 \parallel (r_6 + R_r)$, см. (7-19а)].

Сравнивая (8-19) с (7-15), видим, что входное сопротивление значительно больше, чем в обычном каскаде (поскольку $R_0 \gg r_6$). Это вполне отвечает свойствам последовательной обратной связи. Например, если $R_0 = 500$ Ом и $\beta_{oe} = 50$, то $R_{вх} \approx 20$ кОм.

Заменяв r_6 на R_0 в формуле (8-17в), получим:

$$R_{вых} = r_k^* (1 + \beta \gamma'_6), \quad (8-20)$$

где

$$\gamma'_6 \approx \frac{R_0}{R_r + r_6 + R_0}.$$

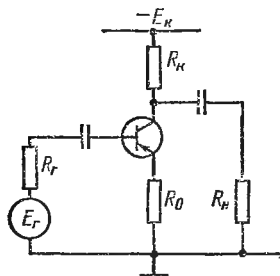


Рис. 8-5. Упрощенная схема каскада с местной последовательной обратной связью по току.

Сравнивая (8-20) с (7-23), приходим к выводу, что выходное сопротивление тоже больше, чем в обычном каскаде (из-за того, что $\gamma'_6 > \gamma_6$). Это обстоятельство соот-

ветствует общим свойствам обратной связи по току. Например, если $R_r = 3R_0 = 1,5$ кОм, $\beta \approx 50$ и $r_k = 1$ МОм, то $R_{вых} \approx 250$ кОм.

Заменяв r_6 на R_0 в формуле (8-17а) и умножив крутизну на суммарную нагрузку $R_k \parallel R_n$, получим коэффициент усиления напряжения:

$$K_u = - \frac{\beta_{oe} (R_k \parallel R_n)}{(R_r + r_6 + R_0) (1 + \beta_{oe} \gamma'_6)} = - \frac{\beta_{oe} (R_k \parallel R_n)}{R_r + R_{вх}}. \quad (8-21)$$

Сравнивая (8-21) с (7-4а) и (7-16), убеждаемся, что коэффициент усиления меньше, чем в обычном каскаде, как и должно быть при отрицательной обратной связи. Если связь глубокая, т. е. $\beta_{oe} \gamma'_6 \gg 1$ (или, что то же самое, $R_r \gg R_{вх}$), то

$$K_u \approx - \frac{R_k \parallel R_n}{R_0}. \quad (8-22а)$$

Отсюда ясно, что значение сопротивления R_0 ограничено условием $R_0 \ll R_k \parallel R_n$ (это условие использовалось выше при записи

β_{oe}).¹ При холостом ходе ($R_n = \infty$) или при работе данного каскада на другой аналогичный каскад ($R_n = R_{вх} \gg R_k$)

$$K_u \approx -\frac{R_k}{R_0}. \quad (8-226)$$

Выражения (8-22) соответствуют общему выражению (8-3), характеризующему глубокую обратную связь, когда усилительный параметр почти не зависит от свойств цепи прямой передачи и определяется коэффициентом обратной связи (в данном случае $\kappa_s = -R_0$).

В области низших частот действительны все формулы § 7-3, за исключением (7-30б), в которой следует заменить r_s на R_0 . Большие значения R_0 и $R_{вх}$ в каскаде с обратной связью позволяют соответственно уменьшить переходные и блокирующие емкости.

В области малых времен и высших частот данный каскад имеет лучшие переходные и частотные характеристики, чем обычный каскад ОЭ. Это следует из выражения (7-41), где β_{oe} и τ_{oe} выражаются формулами (7-20) и (7-36), а коэффициент γ'_6 больше, чем в обычном каскаде ОЭ. При типичном условии $\beta_{oe}\gamma'_6 \gg 1$ получаем постоянную времени в области высших частот:

$$\tau_v = \frac{\tau_{oe}}{1 + \beta_{oe}\gamma'_6} \approx \left(1 + \frac{R_r + r_6}{R_0}\right) [\tau_a + C_k (R_k \parallel R_n)]. \quad (8-23)$$

Добротность каскада, как отмечалось в § 7-4, ниже, чем в отсутствие обратной связи, и для восстановления добротности используется корректирующая емкость C_0 (рис. 7-16).

Местная обратная связь по напряжению. Типичная схема каскада с такой связью показана на рис. 8-6, а. В данном случае, когда каскад является промежуточным, сопротивление обратной связи принято включать между коллекторами; при этом сигнал обратной связи попадает через переходную емкость C_1 на базу этого же транзистора. Такую схему можно использовать и в одиночном каскаде.

Наличие внутренней обратной связи (через r_s) наряду с внешней (через $R_{o.c}$) делает обратную связь в целом многопетлевой. Чтобы не усложнять анализа, будем учитывать влияние только внешней связи; влияние внутренней связи отражается в параметрах цепи прямой передачи. Так как в данном случае обратная связь параллельная по напряжению, то в качестве параметров выбираем входную и выходную проводимости, а также сопротивление передачи (см. табл. 8-1).

Сопротивление передачи — параметр, который в гл. 7 не использовался. Его можно получить либо непосредственно из определения $R_n = (U_{вых}/I_6)_{R_n=\infty}$ (воспользовавшись эквивалентной схемой на рис. 7-4), либо с помощью соотношений, приведенных

¹ Для типичных значений $R_k \parallel R_n = 3 \div 5$ кОм сопротивление R_0 не должно превышать 0,5 кОм, что всегда выполняется.

в табл. 8-2, исходя, например, из параметра K_u . Умножая обе части (7-4а) на R_r , заменяя β на β_{oe} и полагая $R_r, R_n \rightarrow \infty$ (условие, соответствующее идеально обратной связи рассматриваемого типа), получаем:

$$R_n = -\beta (r_k^* \parallel R_k). \quad (8-24)$$

Цепь обратной передачи (резистор $R_{o.c}$ на рис. 8-6, б) будет однонаправленной, если ток $I_{o.c}$ определяется только (или в основном) напряжением $U_{вых}$ (которое является входным для четырехполюсника обратной связи, см. сноску ² на с. 357). Для этого необходимо выполнить неравенство $U_{вх} \ll U_{вых}$. Учитывая соотношения

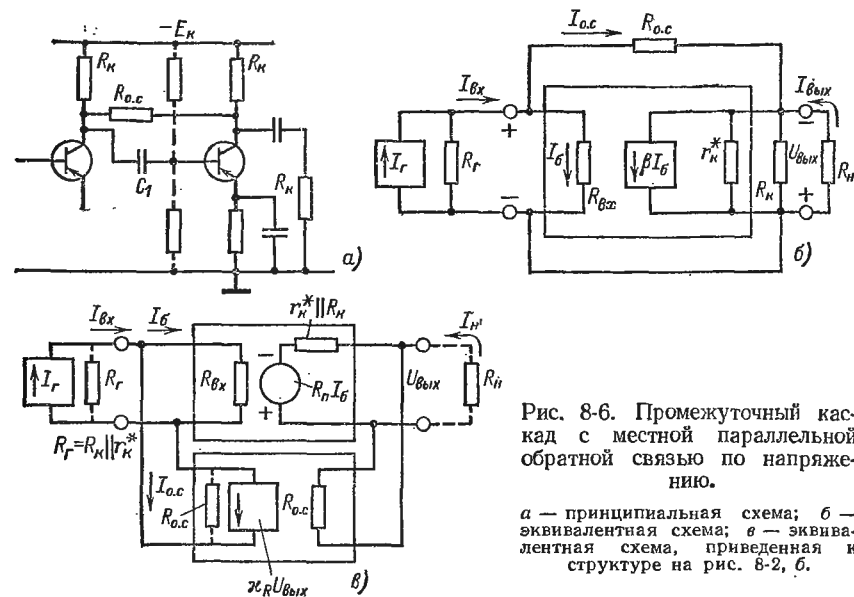


Рис. 8-6. Промежуточный каскад с местной параллельной обратной связью по напряжению.

а — принципиальная схема; б — эквивалентная схема; в — эквивалентная схема, приведенная к структуре на рис. 8-2, б.

$U_{вх} = I_6 R_{вх}$; $U_{вых} \approx \beta I_6 (r_k^* \parallel R_k \parallel R_n \parallel R_{o.c})$ и $R_{вх} \ll \beta (r_k^* \parallel R_k \parallel R_n)$, можно получить условие однонаправленности в следующем виде:

$$R_{o.c} \geq \frac{R_{вх}}{\beta}. \quad (8-25)$$

Поскольку выполнение этого неравенства не встречает затруднений, эквивалентную схему каскада (рис. 8-6, б) можно представить в канонической форме (рис. 8-6, в), которая соответствует общей структуре параллельной обратной связи по напряжению (рис. 8-2, б).

Параметрами однонаправленной цепи обратной связи будем:

$$k_R = -g_{o.c}; \quad (8-26a)$$

$$g_{вхк} = g_{выхк} = g_{o.c}. \quad (8-26б)$$

Как известно, коэффициенты входа и выхода, существенно меньшие единицы, нетипичны и свидетельствуют о неудачно выбранном типе обратной связи (см. с. 363). Поэтому в дальнейшем положим $\xi_{вх} = \xi_{вых} = 1$, т. е. будем считать сопротивления источника сигнала и нагрузки достаточно большими. Тогда глубину обратной связи можно записать в следующем виде:

$$1 + \kappa_R R_n = 1 + \beta \frac{r_k^* \| R_k}{R_{o.c.}},$$

а условие глубокой обратной связи — в виде

$$\beta (r_k^* \| R_k) \gg R_{o.c.} \quad (8-27)$$

Теперь, используя выражения, приведенные в табл. 8-1, нетрудно получить любой из основных параметров каскада с обратной связью. В случае глубокой обратной связи [см. (8-27)] получаем:

$$(R_{вх})_{o.c.} \approx \frac{r_{э} R_{o.c.}}{R_k}; \quad (8-28a)$$

$$(R_{вых})_{o.c.} \approx \frac{R_{o.c.}}{\beta}; \quad (8-28б)$$

$$(R_n)_{o.c.} \approx -R_{o.c.} \quad (8-28в)$$

Запишем еще коэффициент усиления напряжения как более привычный параметр, чем R_n . Для этого, разумеется, нужно считать R_r конечной величиной. Подставим в выражение для $(R_n)_{o.c.}$ (табл. 8-1) значения $(R_n)_{o.c.} = K_u R_r$ и $R_u = K_n R_r$ из табл. 8-2 и учтем соотношение (8-26a).

Тогда в общем случае

$$K_{u.o.c.} = \frac{K_n}{1 - K_n (R_r / R_{o.c.})} \quad (K_u < 0), \quad (8-29a)$$

а при глубокой обратной связи

$$K_{u.o.c.} \approx -\frac{R_{o.c.}}{R_r}. \quad (8-29б)$$

Как видно из (8-29б), с уменьшением сопротивления R_r коэффициент усиления возрастает. Однако одновременно согласно (8-29a) уменьшается глубина обратной связи, т. е. утрачиваются ее положительные свойства. При $R_r \approx 0$ данный тип обратной связи лишен смысла.

Подставляя $R_{o.c.}$ из (8-29б) в (8-27), получаем ограничение на коэффициент усиления в условиях глубокой обратной связи:

$$|K_u| \ll \beta \frac{r_k^* \| R_k}{R_r}. \quad (8-30)$$

Например, если $R_r = r_k^* \| R_k$ (т. е. если выходное сопротивление предыдущего каскада такое же, как у данного), то $|K_u| \ll \beta$, т. е. обычно не превышает 5—10.

Из выражений (8-28а) и (8-28б) следует, что входное и выходное сопротивления под действием обратной связи существенно уменьшаются; это характерно для данного типа обратной связи. Практически при токе $I_0 = 1$ мА входное сопротивление составляет 100—300 Ом, а выходное — около 1 кОм.

Частотные и переходные характеристики коэффициента усиления каскада можно оценить, подставляя изображение (7-39) в (8-29а). После преобразований получим:

$$K_{u.o.c}(s) = \frac{K_{u.o.c}}{1 + s(\tau_v)_{o.c}},$$

где $K_{u.o.c}$ выражается формулой (8-29а), а постоянная времени

$$(\tau_v)_{o.c} = \frac{\tau_v}{1 - K_u R_r / R_{o.c}} \approx \tau_v \frac{R_{o.c}}{-K_u R_r} \quad (K_u < 0). \quad (8-31a)$$

Используя формулы (7-41), (7-36) и (7-16) и полагая $R_r \gg r_0 + r_s$, нетрудно представить постоянную времени в виде

$$(\tau_v)_{o.c} \approx \frac{R_{o.c}}{R_k \parallel R_n} [\tau_a + C_k (R_k \parallel R_n)]. \quad (8-31б)$$

В области низших частот пониженные значения $R_{вх}$ ухудшают частотные свойства и заставляют использовать большие переходные емкости.

Общая обратная связь. *Общая обратная связь выгоднее местной при том же коэффициенте усиления.* Проиллюстрируем это важное положение.

Пусть имеется N -каскадный усилитель с коэффициентом усиления $A_0 = A_1^N$, где A_1 — усиление одного каскада. Рассмотрим два случая:

1) вводится общая обратная связь с глубиной $1 + \kappa_0 A_0$;

2) вводится местная обратная связь в каждом каскаде с глубиной $1 + \kappa_1 A_1$.

Примем, что результирующий коэффициент усиления в обоих случаях одинаков, тогда

$$A_{o.c} = \frac{A_0}{1 + \kappa_0 A_0} = \left(\frac{A_1}{1 + \kappa_1 A_1} \right)^N.$$

Взяв производную $dA_{o.c}/dA_0$, нетрудно получить выражение типа (8-2) для относительной неустойчивости в случае общей обратной связи:

$$\delta_{общ} = \left(\frac{dA_{o.c}}{A_{o.c}} \right)_{общ} = \frac{dA_0}{A_0} \frac{1}{1 + \kappa_0 A_0}.$$

Аналогично, взяв производную $dA_{o.c}/dA_1$, получим относительную неустойчивость при местной обратной связи:

$$\delta_{местн} = \left(\frac{dA_{o.c}}{A_{o.c}} \right)_{местн} = \frac{dA_1}{A_1} \frac{N}{1 + \kappa_1 A_1}.$$

Из исходного равенства $A_0 = A_1^N$ находим: $dA_0/A_0 = N(dA_1/A_1)$, а из выражения для $A_{o.c}$ (см. выше) следует: $1 + \kappa_0 A_0 = (1 + \kappa_1 A_1)^N$. С учетом обоих соотношений запишем искомую связь в следующем виде:

$$\frac{\delta_{общ}}{\delta_{местн}} = \frac{1}{(1 + \kappa_1 A_1)^{N-1}} = \frac{1}{(1 + \kappa_0 A_0)^{\frac{N-1}{N}}}.$$

Отсюда ясно, что $\delta_{\text{общ}} < \delta_{\text{местн}}$, причем для г л у б о к ой обратной связи разница может быть весьма значительной. Например, если $N = 2$ и $1 + \chi_0 A_0 = 25$, то $\delta_{\text{общ}} = 0,28 \delta_{\text{местн}}$. С ростом числа каскадов выигрыш увеличивается.

Как известно [62], использование общей обратной связи в усилителях с числом каскадов больше трех обычно приводит к самовозбуждению или во всяком случае должно сопровождаться специальными мерами стабилизации. Поэтому рассмотрим только случаи двух- и трехкаскадных усилителей.

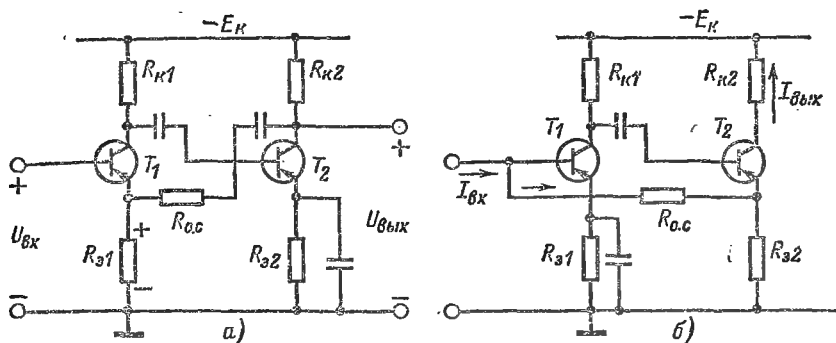


Рис. 8-7. Двухкаскадные усилители с общей обратной связью.
 а — последовательной по напряжению; б — параллельной по току.

Схема обратной связи, как обычно, выбирается, исходя из назначения усилителя, т. е. в зависимости от характера входной и выходной величин (ток или напряжение), а также от желательного изменения в ту или иную сторону входного и выходного сопротивления.

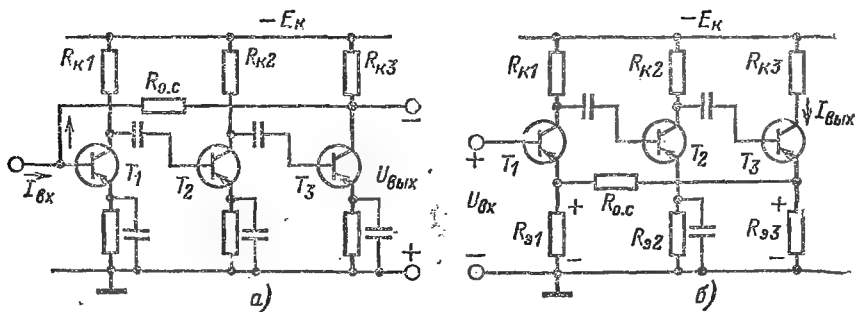


Рис. 8-8. Трехкаскадные усилители с общей обратной связью.
 а — параллельной по напряжению; б — последовательной по току.

ний. На рис. 8-7 показаны два типичных варианта обратной связи в двухкаскадных усилителях, а на рис. 8-8 — в трехкаскадных.

Если сравнить анализ такого рода схем с анализом рассмотренных выше однокаскадных усилителей, то окажется, что он сложнее в части переходных и частотных характеристик, а также вопросов стабильности, но проще в области средних частот. Последнее объясняется тем, что в многокаскадных усилителях цепь прямой

передачи характеризуется таким высоким значением коэффициента передачи A , при котором глубокая обратная связь обеспечивается при весьма малых значениях коэффициента обратной связи κ . Малые значения κ способствуют однонаправленности цепи обратной связи, а также слабому влиянию ее входных и выходных параметров на коэффициенты $\xi_{вх}$ и $\xi_{вых}$. Кроме того, большие значения A позволяют считать практически неизменной входную величину цепи прямой передачи (см. m_1 на рис. 8-1).

Все эти особенности существенно облегчают расчеты, тем более, что в области средних частот коэффициент передачи при глубокой обратной связи выражается элементарной формулой (8-3).

В схеме на рис. 8-7, а (которую часто называют «двойкой») действует последовательная обратная связь по напряжению¹. Значит, усильтельным параметром согласно рис. 8-2, а целесообразно считать коэффициент усиления K_u .

В общем случае, как следует из табл. 8-1,

$$K_{u.o.c} = \frac{K_u}{1 + \kappa_u K_u}. \quad (8-32a)$$

Коэффициент K_u рассчитывается по формуле (7-58a). Полагая $\beta_{oe} = \beta$; $R_r \ll R_{вх1} \approx \beta_1 R_{э1}$ и $R_{к1} \gg R_{вх2}$, получаем:

$$K_u \approx \beta_2 \frac{R_{к2}}{R_{э1}}.$$

Если $\Delta I_{э1} \ll \Delta I_{o.c}$ (что обычно выполняется), то κ_u есть коэффициент передачи делителя $R_{э1} - R_{o.c}$, т. е.

$$\kappa_u \approx \frac{R_{э1}}{R_{э1} + R_{o.c}}.$$

Следовательно, при глубокой обратной связи выражение (8-32a) переходит в следующее:

$$K_{u.o.c} \approx \frac{1}{\kappa_u} = 1 + \frac{R_{o.c}}{R_{э1}} \approx \frac{R_{o.c}}{R_{э1}}. \quad (8-32б)$$

Условие $R_{o.c} \gg R_{э1}$, при котором $K_{u.o.c} \gg 1$, легко выполнимо: оно не противоречит условию глубокой обратной связи ($\kappa_u K_u \gg 1$) благодаря большому усилению K_u . Типичными значениями являются: $K_u = 500 \div 2000$; $R_{o.c} = 5 \div 10$ кОм; $R_{э1} = 50 \div 200$ Ом; при этом $1 + \kappa_u K_u = 20 \div 50$ и $K_{u.o.c} = 25 \div 40$.

Глубокая обратная связь и высокая устойчивость делают данную схему удобным элементом *линейных импульсных усилителей* [123].

¹ Строго говоря, в схеме действует двухпетлевая обратная связь: общая по напряжению и местная (в 1-м каскаде) по току. Если представить выход цепи так, как показано на рис. 8-2, а, то резистор $R_{э1}$ можно отнести к цепи прямой передачи. Тогда усиление и входное сопротивление 1-го каскада можно рассчитывать по формулам (8-21) и (8-19).

Если считать, что в отсутствие обратной связи $R_{\text{вых}} = R_{\kappa 2}$ и $R_{\text{вх}} \approx \beta_1 R_{\text{эл}}$, то при глубокой обратной связи входное и выходное сопротивления схемы выражаются следующим образом:

$$R_{\text{вх. о. с}} = R_{\text{вх}} (1 + \kappa_n K_n) \approx \beta_1 \beta_2 \frac{R_{\text{эл}} R_{\kappa 2}}{R_{\text{о. с}}}; \quad (8-33a)$$

$$R_{\text{вых. о. с}} = \frac{R_{\text{вых}}}{1 + \kappa_n K_n} \approx \frac{R_{\text{о. с}}}{\beta_2}, \quad (8-33b)$$

то есть входное сопротивление повышено, а выходное понижено. Типичными значениями будут: $R_{\text{вх. о. с}} = 20 \div 50$ кОм; $R_{\text{вых. о. с}} = 100 \div 200$ Ом.

Часто в схему вводят третий каскад — эмиттерный повторитель (см. гл. 9) — и обратную связь подают с выхода повторителя. Такая модификация (так называемая «тройка») не требует принципиальных изменений в анализе и его результатах, поскольку выходное напряжение повторителя практически совпадает с напряжением $U_{\kappa 2}$ по величине и фазе¹.

Преимуществами при введении третьего каскада являются гораздо меньшее выходное сопротивление (до 10 Ом и меньше) и соответственно возможность уменьшения сопротивления $R_{\text{о. с}}$, а вместе с ним $R_{\text{эл}}$. Уменьшение последнего ослабляет местную обратную связь в 1-м каскаде (см. сноску на с. 376) и тем самым повышает исходный коэффициент усиления K_n , а следовательно, и глубину общей обратной связи в несколько раз.

В схеме на рис. 8-7, б действует параллельная обратная связь по току и, следовательно, усилительным параметром согласно рис. 8-2, г целесообразно считать K_i . В общем случае, как следует из табл. 8-1,

$$K_{i \text{ о. с}} = \frac{K_i}{1 + \kappa_i K_i}. \quad (8-34a)$$

Полагая $\beta_{\text{о. с}} = \beta$, $R_{\kappa 1} \gg R_{\text{вх} 2}$ и $R_{\text{г}} \gg R_{\text{вх} 1}$, можно считать $K_i \approx \beta_1 \beta_2$.

Если $\Delta U_{\text{б} 1} \ll \Delta U_{\text{эл} 2}$ и $\Delta I_{\text{вых}} \approx \Delta I_{\text{эл} 2} \gg \Delta I_{\text{о. с}}$, то $\Delta I_{\text{о. с}} \approx (\Delta I_{\text{вых}} R_{\text{эл} 2}) / R_{\text{о. с}}$ и тогда

$$\kappa_i = \Delta I_{\text{о. с}} / \Delta I_{\text{вых}} \approx R_{\text{эл} 2} / R_{\text{о. с}}.$$

При условии глубокой обратной связи выражение (8-34a) переходит в следующее:

$$K_{i \text{ о. с}} \approx \frac{1}{\kappa_i} \approx \frac{R_{\text{о. с}}}{R_{\text{эл} 2}}. \quad (8-34b)$$

Условие $R_{\text{о. с}} \gg R_{\text{эл} 2}$, при котором $K_{i \text{ о. с}} \gg 1$, легко выполняется; оно не противоречит условию глубокой обратной связи ($\kappa_i K_i \gg 1$) благодаря большому усилению K_i . Типичными значениями являются: $K_i = 500 \div 2000$; $R_{\text{о. с}} = 0,5 \div 2$ кОм; $R_{\text{эл} 2} = 50 \div 200$ Ом;

¹ Исключение, как и в ламповом варианте [62], составляет область малых времен и высоких частот, когда дополнительный сдвиг фазы в повторителе может понизить устойчивость усилителя или вызвать самовозбуждение.

при этом $1 + \kappa_i K_i = 10 \div 40$ и $K_{i.o.c} = 20 \div 50$. Считая, что в отсутствие обратной связи $R_{вх} \approx \beta_1 r_{э1}$ и $R_{вых} \approx r_{кз}^*$, для входного и выходного сопротивлений (при глубокой обратной связи) получаем:

$$R_{вх.o.c} = \frac{R_{вх}}{1 + \kappa_i K_i} \approx \frac{r_{э1} R_{o.c}}{\beta_1 R_{э2}}; \quad (8-35a)$$

$$R_{вых.o.c} = R_{вых} (1 + \kappa_i K_i) \approx \frac{\beta_1 r_{к2} R_{э2}}{R_{o.c}}. \quad (8-35б)$$

Очевидно, что входное сопротивление понижено, а выходное — повышено. Типичными значениями будут: $R_{вх.o.c} = 10 \div 50 \text{ Ом}$; $R_{вых.o.c} = 0,2 \div 1 \text{ МОм}$.

Такие усилители удобны, например, в ядерной электронике при работе от фотоэлектронных умножителей (ФЭУ), имеющих очень большое внутреннее сопротивление [124]. Интересно отметить, что, включая на входе последовательное сопротивление $r_1 \gg R_{вх.o.c}$, а на выходе параллельное сопротивление $r_2 \ll R_{вых.o.c}$, получаем по-прежнему усилитель тока, но с входным и выходным сопротивлениями, близкими к r_1 и r_2 , т. е. практически не зависящими от режима [121].

В схеме на рис. 8-8, а действует параллельная обратная связь по напряжению и, следовательно, усилительным параметром согласно рис. 8-2, б целесообразно считать сопротивление передачи $R_{п}$. Данная схема является обобщением однокаскадной (см. рис. 8-6). При условии $\Delta U_{кз} \gg \Delta U_{б1}$, которое всегда выполняется, имеем:

$$\kappa_R = \frac{\Delta I_{o.c}}{\Delta U_{вых}} \approx -\frac{1}{R_{o.c}}.$$

Следовательно, в общем виде

$$R_{п.o.c} = \frac{R_{п}}{1 + \kappa_R R_{п}}, \quad (8-36a)$$

а при глубокой обратной связи

$$R_{п.o.c} \approx -R_{o.c}. \quad (8-36б)$$

В отличие от однокаскадной схемы условие глубокой связи $-R_{п}/R_{o.c} \gg 1$ легко выполнимо, поскольку исходное сопротивление передачи имеет очень большую величину:

$$R_{п} = \frac{U_{вых}}{I_{вх}} = -\frac{I_{кз} R_{кз}}{I_{б1}} \approx -\beta_1 \beta_2 \beta_3 R_{кз}$$

(обычно 50—300 МОм и больше). Поэтому даже при условии $R_{o.c} \gg R_{кз}$, при котором $R_{o.c} = 10 \div 100 \text{ кОм}$, обеспечиваются глубина обратной связи в несколько тысяч и весьма высокая стабильность.

Входное и выходное сопротивления, которые в отсутствие обратной связи равны $R_{вх} \approx \beta_1 r_{э1}$ и $R_{вых} \approx R_{кз}$, в рабочей схеме

равны:

$$R_{\text{вх. о. с}} = \frac{R_{\text{вх}}}{1 + \kappa_R R_{\Pi}} \approx r_{\text{э1}} \frac{R_{\text{о. с}}}{\beta_2 \beta_3 R_{\text{кз}}}; \quad (8-37a)$$

$$R_{\text{вых. о. с}} = \frac{R_{\text{вых}}}{1 + \kappa_R R_{\Pi}} \approx \frac{R_{\text{о. с}}}{\beta_1 \beta_2 \beta_3}. \quad (8-37b)$$

Оба сопротивления оказываются очень малыми: $R_{\text{вх. о. с}} = 0,1 \div 1 \text{ Ом}$; $R_{\text{вых}} = 0,5 \div 5 \text{ Ом}$. Такой усилитель, как и предыдущий, удобен для работы с генераторами тока, например ФЭУ.

Зависимость входного сопротивления от тока $I_{\text{э1}}$ можно предотвратить последовательным включением сопротивления $r_1 \gg R_{\text{вх. о. с}}$, как указывалось выше. Если вместо R_{Π} использовать параметр K_{Π} , то согласно табл. 8-2 и формуле (8-36b) при глубокой обратной связи получаем $K_{\Pi} \approx -R_{\text{о. с}}/R_{\Gamma}$, т. е. коэффициент усиления зависит от сопротивления источника сигнала. При $R_{\Gamma} = 10 \div 100 \text{ Ом}$ значение K_{Π} может составлять несколько тысяч.

В схеме на рис. 8-8, б действует последовательная обратная связь по току и усилительным параметром согласно рис. 8-2, в следует считать крутизну. Если $\Delta U_{\text{э1}} \ll \Delta U_{\text{э3}}$, что всегда легко обеспечить, то

$$\kappa_S = \frac{\Delta U_{\text{о. с}}}{\Delta I_{\text{вых}}} \approx - \frac{R_{\text{э1}} R_{\text{э3}}}{R_{\text{о. с}}}.$$

Тогда в общем случае

$$S_{y. \text{ о. с}} = \frac{S_y}{1 + \kappa_S S_y}, \quad (8-38a)$$

а при глубокой обратной связи

$$S_{y. \text{ о. с}} \approx - \frac{R_{\text{о. с}}}{R_{\text{э1}} R_{\text{э3}}}. \quad (8-38b)$$

Если принять $\beta_{\text{оэ}} = \beta$; $R_{\text{вх1}} = \beta_1 R_{\text{э1}}$ и $R_{\text{к2}} \gg R_{\text{вх3}}$, то в отсутствие обратной связи крутизна S_y выражается формулой

$$S_y = \frac{I_{\text{вых}}}{U_{\text{вх}}} = - \frac{I_{\text{к3}}}{I_{\text{б1}} R_{\text{вх1}}} \approx - \frac{\beta_2 \beta_3}{R_{\text{э1}}}$$

и может составить 10—15 А/В.

Типичными параметрами схемы являются: $R_{\text{э1}} = 50 \div 200 \text{ Ом}$; $R_{\text{э3}} = 10 \div 20 \text{ Ом}$; $R_{\text{о. с}} = 0,5 \div 2 \text{ кОм}$; при этом получается $1 + \kappa_S S_y = 20 \div 50$ и $S_{y. \text{ о. с}} = 0,5 \div 2 \text{ А/В}$.

Считая, что в отсутствие обратной связи $R_{\text{вх}} \approx \beta_1 R_{\text{э1}}$ и $R_{\text{вых}} \approx r_{\text{кз}}^*$, получаем для входного и выходного сопротивлений рабочей схемы:

$$R_{\text{вх. о. с}} = R_{\text{вх}} (1 + \kappa_R S_y) \approx \beta_1 \beta_2 \beta_3 \frac{R_{\text{э1}} R_{\text{э3}}}{R_{\text{о. с}}}; \quad (8-39a)$$

$$R_{\text{вых. о. с}} = R_{\text{вых}} (1 + \kappa_S S_y) \approx \beta_2 \frac{r_{\text{кз}} R_{\text{э3}}}{R_{\text{о. с}}}. \quad (8-39b)$$

Оба сопротивления имеют большие расчетные значения, например $R_{\text{вх. о. с}} = 200 \div 500 \text{ кОм}$; $R_{\text{вых. о. с}} = 0,5 \div 1 \text{ МОм}$. Фактически

предельными значениями того и другого являются соответственно сопротивлению $r_{к1}$ и $r_{к3}$, которые не учитывались при записи коэффициентов β .

Вариантами рассмотренной схемы могут быть усилители, у которых $R_{о.с} = 0$, а сопротивление R_9 является общим для 1-го и 3-го или для всех трех каскадов. В этих случаях обратная связь по току является 100%-ной, т. е.

$$\kappa_S = \frac{\Delta U_{о.с}}{-\Delta I_{вых}} \approx -\frac{\Delta I_{93} R_9}{\Delta I_{93}} = -R_9$$

и соответственно

$$S_{y. о. с} \approx -\frac{1}{R_9}. \quad (8-40a)$$

Чтобы крутизна была того же порядка, что и в схеме на рис. 8-9, б (т. е. около 1 А/В), необходимы значения $R_9 = 0,5 \div 2$ Ом. При этом в отсутствие обратной связи $R_{вх} \approx \beta_1 r_{91}$ и $S_y \approx -\beta_2 \beta_3 / r_{91}$, а входное и выходное сопротивления рабочей схемы имеют значения:

$$R_{вх. о. с} \approx \beta_1 \beta_2 \beta_3 R_9; \quad (8-40б)$$

$$R_{вых. о. с} = \beta_2 r_{к3} \frac{R_9}{r_{91}}. \quad (8-40в)$$

Порядок величин этих сопротивлений тот же, что и в основной схеме (независимо от расчетных значений, они тоже ограничены значениями $r_{к1}$ и $r_{к3}$).

8-5. ОБРАТНАЯ СВЯЗЬ ПО ПОСТОЯННОМУ ТОКУ

С обратной связью по постоянному току мы познакомились в гл. 6 при анализе режима и стабильности рабочей точки. В частности, выражения (6-4) и (6-8) содержат характерный для внутренней обратной связи двучлен $1 + \beta \gamma_6$ — глубину обратной связи, где коэффициент обратной связи γ_6 определяется внешними сопротивлениями схемы.

Теперь рассмотрим пример более сложной обратной связи по постоянному току, назначение которой, однако, такое же, как и в отдельном каскаде: стабилизация режима транзисторов по отношению к температуре, допускам и другим факторам.

На рис. 8-9, а показана распространенная в свое время схема трехкаскадного усилителя с общей и местными обратными связями по постоянному току [125]. Обратная связь по переменному току предотвращается с помощью конденсатора C большой емкости, чтобы не снижать усиления сигнала. Такая схема, в основе которой лежит усилитель с гальванической связью каскадов, позволяет существенно улучшить передачу низших частот или вершины импульсов, так как в ней резко уменьшено количество переходных и блокирующих конденсаторов. Отсутствие последних приводит, разумеется, к снижению усиления в цепи прямой передачи, но наличие трех каскадов обычно компенсирует этот недостаток.

По отношению к п р и р а щ е н и я м постоянных составляющих в схеме действует параллельная отрицательная обратная связь по напряжению. Такая связь стабилизирует коллекторный потенциал транзистора T_3 , а следовательно, и его коллекторный ток. Для того чтобы оценить степень стабилизации, нужно найти глубину обратной связи.

При достаточно большом усилении ($K_u > 10$) и достаточно большой величине сопротивления обратной связи ($R_{о.с} \gg R_K$) можно считать цепь обратной связи однонаправленной и изобразить усилитель в виде скелетной схемы на рис. 8-9, б.

Эта схема соответствует структуре на рис. 8-2, б. Отсюда следует, что усилительным параметром является сопротивление передачи. Коэффициент обратной связи $K_R \approx -1/R_{o.c.}$, а глубина обратной связи, определяющая стабильность, составляет $1 - R_{\Pi}/R_{o.c.}$.

Поскольку сопротивление передачи редко используется как параметр, его можно выразить через более распространенную величину $K_i = R_{\Pi}/R_{\Pi}$ (см. табл. 8-2). Тогда получаем глубину обратной связи $1 - K_i R_{\Pi}/R_{o.c.}$, где под R_{Π} следует понимать сопротивление $R_{кз}$. Например, если $R_{кз} = 5$ кОм, $R_{o.c.} = 1$ МОм и $K_i = -5000$, то $1 + \mu A = 25^*$.

Общая обратная связь позволяет уменьшить результирующий дрейф тем сильнее, чем больше параметр A . Поэтому часто имеет смысл уменьшить сопротивления R_{Π} (или даже заменить их полупроводниковыми стабилизаторами).

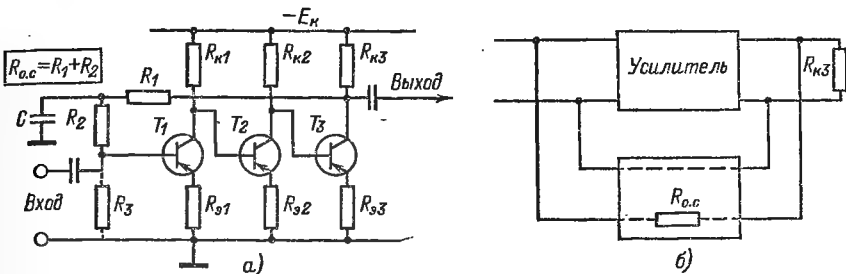


Рис. 8-9. Трехкаскадный усилитель, стабилизированный общей параллельной обратной связью по напряжению постоянного тока.

а — принципиальная схема; б — скелетная схема.

Получающееся при этом возрастание коэффициентов неустойчивости S в каждом из каскадов (см. § 6-2) с избытком компенсируется возрастанием коэффициента усиления K_i и, следовательно, глубины обратной связи. Этот вывод соответствует общему положению о предпочтительности общей обратной связи перед местными (см. § 8-4).

Рассмотренный метод стабилизации может иметь ряд вариантов. Например, в том же трехкаскадном усилителе можно подавать обратную связь с эмиттера транзистора T_3 на эмиттер транзистора T_1 . В двухкаскадном усилителе можно подавать обратную связь с коллектора транзистора T_2 на эмиттер транзистора T_1 и т. д.

Иногда на входе добавляют резистор R_3 (на рис. 8-9, а показан пунктиром), который уменьшает результирующее сопротивление в цепи базы транзистора T_1 и тем самым его коэффициент неустойчивости (см. § 6-2). Такой метод стабилизации можно считать развитием метода, используемого иногда в отдельных каскадах (см. рис. 6-6), или, наоборот, последний можно считать частным случаем общего метода.

Глава девятая

ЭМИТТЕРНЫЕ ПОВТОРИТЕЛИ

9-1. ВВЕДЕНИЕ

Часто возникает необходимость в каскадах с повышенным входным сопротивлением и пониженным выходным сопротивлением.

Наиболее распространенным каскадом такого типа является эмиттерный повторитель — схемный аналог катодного повторителя

* Сопротивление $R_{o.c.}$ в данной схеме выбирается из очевидного условия $R_{o.c.} = (U_{кз} - U_{б1})/I_{б1}$, где $I_{б1}$ — желательный ток базы. В приведенном примере положено $I_{б1} = 10$ мкА и $U_{кз} - U_{б1} = 10$ В.

162]. Эмиттерный повторитель (рис. 9-1, а) обеспечивает включение транзистора по схеме ОК; он не дает усиления по напряжению и не поворачивает фазу входного сигнала. В этом легко убедиться, если

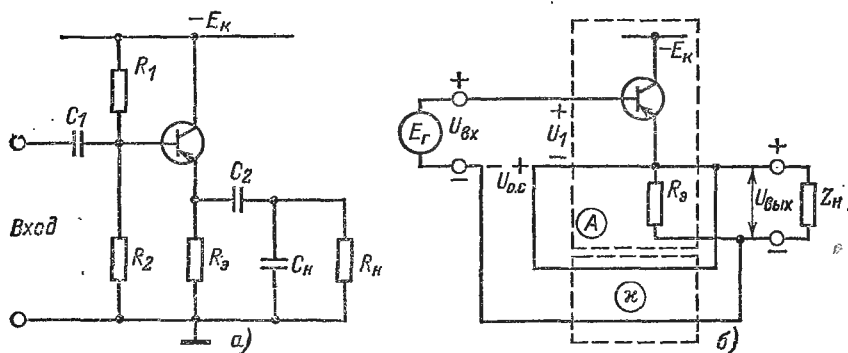


Рис. 9-1. Принципиальная схема простого повторителя.

а — в обычном изображении; б — приведенная к структуре на рис. 8-9, а.

учесть, что напряжение $U_{бэ}$ мало меняется при работе каскада. Следовательно, потенциал эмиттера «привязан» к потенциалу базы и повторяет его изменения.

9-2. ПРОСТЫЙ ПОВТОРИТЕЛЬ

Схему, показанную на рис. 9-1, а называют простым повторителем в отличие от сложных, которые содержат два транзистора и более и будут рассмотрены позднее. Анализируя повторители, мы не будем затрагивать область низших частот, которая не имеет специфики по сравнению с каскадом ОЭ. Типичной нагрузкой повторителя будем считать совокупность активного сопротивления и емкости, включенных параллельно:

$$Z_{н.э}(s) = \frac{R_H \parallel R_3}{1 + s\tau_{н.э}},$$

где

$$\tau_{н.э} = C_H (R_H \parallel R_3).$$

Схема на рис. 9-1, а является типичным примером усилителя со 100%-ной последовательной обратной связью по напряжению (рис. 9-1, б). Поэтому ее можно было бы анализировать, исходя из общих соотношений, приведенных в § 8-2. Однако полезнее анализировать схему в целом, не выделяя цепь обратной передачи.

Входное сопротивление. На эквивалентной схеме повторителя (рис. 9-2) заменим комбинацию сопротивлений, расположенных справа от точек $B'-K$, одним сопротивлением $r_k^* \parallel (r_3 + R_3 \parallel R_H)$. Тогда структура данной схемы будет подобна структуре схемы ОЭ (см. рис. 7-2), если в последней положить $R_K = 0$. Поэтому входное

сопротивление повторителя определяется формулой (7-2), в которой сопротивление r_b следует заменить указанным эквивалентным сопротивлением¹:

$$R_{вх} = r_b + (1 + \beta) [r_k^* \| (r_a + R_a \| R_H)]. \quad (9-1a)$$

На практике всегда выполняется соотношение $r_a \ll R_a \| R_H$ (поэтому на рис. 9-2 сопротивление r_a показано пунктиром). В этом случае формула (9-1a) упрощается и становится более наглядной:

$$R_{вх} \approx r_b + (1 + \beta) (r_k^* \| R_a \| R_H). \quad (9-1б)$$

Здесь первое слагаемое обычно пренебрежимо мало по сравнению со вторым. Кроме того, в большинстве случаев внешние сопротивления R_a и R_H значительно меньше, чем r_k^* , и последнее можно не учитывать. Однако оно накладывает принципиальное ограничение на входное сопротивление. Действительно, если предположить $R_a \| R_H \gg r_k^*$, то из (9-1) получим максимально возможное значение входного сопротивления:

$$R_{вх. макс} \approx r_k^*. \quad (9-2)$$

Как известно, сопротивление r_k у маломощных транзисторов доходит до 1 МОм и выше².

Практически при холостом ходе ($R_H = \infty$) и при сопротивлении R_a порядка 10 кОм можно получить $R_{вх}$ до 100—200 кОм. При наличии нагрузки входное сопротивление обычно определяется значением R_H , которая в случае повторителя не бывает большой (иначе не имеет смысла использовать повторитель). Если нагрузка не остается постоянной, то меняется и входное сопротивление.

Величина входного сопротивления сильно ограничивается делителем в цепи базы. Для хорошей стабильности согласно (6-9) желательно иметь $R_1 \| R_2 < R_a$, а для сохранения большого входного сопротивления желательно иметь $R_1 \| R_2 > R_{вх} \approx \beta (R_a \| R_H)$. Сочетание этих условий возможно лишь тогда, когда $R_a > \beta R_H$. Это в свою очередь означает сравнительно низкоомную нагрузку R_H (200—300 Ом), так как R_a не может быть очень большим из-за трудностей в установлении необходимого режима транзистора.

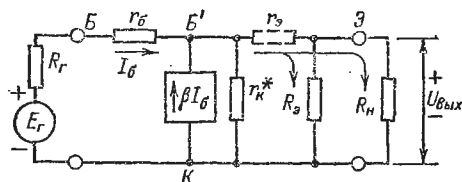


Рис. 9-2. Эквивалентная схема повторителя в области средних частот.

¹ На схеме для простоты не показаны сопротивления делителя R_1 , R_2 . Влияние делителя можно учесть по теореме об эквивалентном генераторе (см. § 7-2).

² При малых рабочих токах (десяти доли миллиампера и менее) можно согласно (4-24) увеличить r_k , а вместе с ним и $R_{вх. макс}$ во много раз. Однако при этом ухудшаются другие параметры повторителя, прежде всего нагрузочная способность.

В схемах, в которых входное сопротивление должно иметь возможно большую величину, приходится либо использовать непосредственную связь с источником сигнала (без делителя), либо искусственно повышать сопротивление цепи смещения (см. § 9-3).

Важнейшая особенность эмиттерного повторителя состоит в том, что его входное сопротивление резко уменьшается в случае коротких импульсов или при повышенной частоте, даже сравнительно небольшой. Такая зависимость, как и в других случаях, обусловлена инерционностью процессов в базе транзистора, а также наличием коллекторной и нагрузочной емкостей. Из рис. 9-3 легко

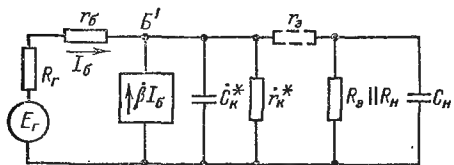


Рис. 9-3. Эквивалентная схема повторителя в области малых времен.

видеть, что в первый момент после подачи импульса входное сопротивление будет равно r_b , т. е. очень мало по сравнению с (9-1). За время переходного процесса $R_{вх}$ увеличивается до значения (9-1).

Перейдем к количественному анализу. Подставив $\beta(s)$, $Z_k^*(s)$ и $Z_{н.э}(s)$ в формулу (9-16), после некоторых преобразований можно привести изображение входного сопротивления к следующему общему виду:

$$Z_{вх}(s) \approx r_b + (1 + \beta) (r_k^* \parallel R_c \parallel R_n) \frac{1 + s\tau_a}{1 + sa_1 + s^2a_2}. \quad (9-3)$$

Здесь

$$a_1 = \tau'_\beta + \tau'_k + \tau'_{н.э};$$

$$a_2 = \tau_\beta \tau'_{н.э} + \tau_a \tau'_k;$$

$$\tau'_\beta = \tau_\beta \gamma_k^*, \quad \tau'_{н.э} = \tau_{н.э} \gamma_k^*, \quad \tau'_k = \tau_k (1 - \gamma_k^*) = C_k^* (r_k^* \parallel R_c \parallel R_n);$$

$$\gamma_k^* = \frac{r_k^*}{r_k^* + R_c \parallel R_n}.$$

Выражение (9-3) подтверждает резкое уменьшение входного сопротивления при $t = 0$ (т. е. при $s = \infty$), а также показывает возможность немонотонного изменения функции $Z_{вх}(t)$. В случае $C_n = 0$ выражение (9-3) упрощается и с некоторым приближением можно представить его в следующем виде:

$$Z_{вх}(s) \approx r_b + \frac{(1 + \beta) (r_k^* \parallel R_c \parallel R_n)}{1 + s(\tau'_\beta + \tau'_k)}. \quad (9-4)$$

Изображению (9-4) соответствует функция $z_{вх}(t)$, которая равна r_b при $t = 0$ и экспоненциально (с постоянной времени $\tau'_\beta + \tau'_k$) возрастает до значения (9-16). Отсюда следует, что для коротких импульсов, с длительностью $t_n < (\tau'_\beta + \tau'_k)$, входное сопротивление

ние существенно меньше статического сопротивления r_6 . С учетом нагрузочной емкости критерием короткого сигнала будет $t_n < (\tau'_\beta + \tau'_k + \tau'_{н.э})$.

В случае бездрейфовых сплавных транзисторов обычно можно считать короткими импульсы с длительностью менее 10 мкс, а в случае дрейфовых — с длительностью менее 1 мкс.

Частотные характеристики $Z_{вх}(\omega)$ получаются заменой оператора s оператором $j\omega$ в формулах (9-3) и (9-4) и имеют, как правило, спадающий характер. Так, при $C_n = 0$ [см. формулу (9-4)] частотная характеристика близка к простейшей функции

$$Z_{вх}(\omega) \approx \frac{R_{вх}}{1 + j \frac{\omega}{\omega_b}}, \quad (9-5)$$

где граничная частота $\omega_b = 1/(\tau'_\beta + \tau'_k)$. Однако при $\omega \rightarrow \infty$ входное сопротивление приближается не к нулю, а к величине r_6 . Для бездрейфовых транзисторов частота f_b обычно лежит в пределах 50—100 кГц; для дрейфовых транзисторов эти пределы примерно на порядок больше.

Кривые, изображающие переходные и частотные характеристики $Z_{вх}$, подобны кривым коэффициента передачи (см. ниже).

Выходное сопротивление. Этот параметр определяет нагрузочную способность повторителя как для активной, так и для комплексной нагрузки. Выходное сопротивление можно найти из рис. 9-2 при $E_r = 0$. Для этого сначала зададим э. д. с. E несколько правее точек B' и K . Тогда $I_6 = E/(R_r + r_6)$, а $I_E = I_6(1 + \beta)$, откуда следует:

$$\frac{E}{I_E} = R_{B'-K} = \frac{R_r + r_6}{1 + \beta}.$$

Теперь, «перемещаясь» от точек $B'-K$ к зажимам $\mathcal{E}-K$, легко записать выходное сопротивление в следующем общем виде:

$$R_{вых} = \left(\frac{R_r + r_6}{1 + \beta} \parallel r_k^* + r_9 \right) \parallel R_9. \quad (9-6a)$$

Обычно сопротивления r_k^* и R_9 в этом выражении малосущественны, и его можно представить в таком виде:

$$R_{вых} \approx r_9 + \frac{R_r + r_6}{1 + \beta}. \quad (9-6б)$$

В случае весьма низкоомного источника сигнала ($R_r < r_6$) и достаточно большой величины β можно иногда пренебречь вторым членом в (9-6б) и получить:

$$R_{вых} \approx r_9. \quad (9-6в)$$

Однако это простейшее соотношение утрачивает силу при больших токах, когда сопротивление r_9 уменьшается.

При токе 1 мА сопротивление $R_{\text{вых}}$ согласно (9-6в) составляет всего 25 Ом и легко может быть уменьшено путем увеличения рабочего тока в допустимых пределах. Минимальное выходное сопротивление (при $r_{\text{э}} \approx 0$ и $R_{\text{г}} = 0$) выражается согласно (9-6б) следующим образом:

$$R_{\text{вых. мин}} = \frac{r_6}{1 + \beta} \quad (9-7a)$$

и может составлять 0,5—2 Ом и меньше¹.

Важно отметить, что в эмиттерном повторителе выходное сопротивление зависит от сопротивления источника сигнала $R_{\text{г}}$. Последнее при малых значениях $r_{\text{э}}$ (например, у мощных транзисторов) может играть определяющую роль. В пределе, если $R_{\text{г}} \rightarrow \infty$, выходное сопротивление согласно (9-6а) стремится к максимальному значению

$$R_{\text{вых. макс}} \approx r_{\text{к}}^* \| R_{\text{э}}. \quad (9-7б)$$

Практически в большинстве случаев $R_{\text{г}} < 2 \div 5$ кОм и тогда $R_{\text{вых}}$ не превышает 100—200 Ом, т. е. намного меньше, чем в каскаде ОЭ.

Иногда нагрузка повторителя меняется скачком. Тогда сопротивление $R_{\text{вых}}$ следует рассматривать как функцию времени. Изображение этой функции получится, если подставить $\beta(s)$ и $Z_{\text{к}}^*(s)$ в формулы (9-6). Не вдаваясь в подробности анализа, заметим следующее. Если бы коллекторная емкость $C_{\text{к}}$ отсутствовала, то в первый момент после скачка нагрузки, когда $\beta = 0$, выходное сопротивление имело бы значение

$$[(R_{\text{г}} + r_6) \| r_{\text{к}} + r_{\text{э}}] \| R_{\text{э}} \approx (R_{\text{г}} + r_6) \| R_{\text{э}},$$

которая может быть гораздо больше, чем установившееся значение (9-6б). При наличии $C_{\text{к}}$ получаем в первый момент $Z_{\text{к}}^*(0) = 0$ и выходное сопротивление определяется формулой (9-6в), которая меньше установившегося значения (9-6б). Однако последующий переходный процесс обычно оказывается немонотонным и сопротивление $R_{\text{вых}}$ сначала всегда возрастает, а затем уменьшается до установившегося значения. Такое временное увеличение выходного сопротивления существенно в случаях импульсной нагрузки повторителя. Анализ показывает, что длительность переходного процесса определяется постоянной времени, близкой к $\tau_{\text{б}}/(1 + \beta\gamma_6)$, где γ_6 имеет величину, приведенную на с. 388.

Коэффициент передачи. Если из рис. 9-2 найти выходное напряжение и положить $r_{\text{э}} \ll r_{\text{к}}^*$, что всегда выполняется, то

$$U_{\text{вых}} = (1 + \beta) I_6 (R_{\text{э}} \| R_{\text{н}} \| r_{\text{к}}^*);$$

$$I_6 = \frac{E_{\text{г}}}{R_{\text{г}} + R_{\text{вх}}}.$$

Отсюда, используя (9-1а), получаем коэффициент передачи напряжения:

$$K_u = \frac{(1 + \beta) (r_{\text{к}}^* \| R_{\text{э}} \| R_{\text{н}})}{R_{\text{г}} + r_6 + (1 + \beta) [r_{\text{к}}^* \| (r_{\text{э}} + R_{\text{э}} \| R_{\text{н}})]}. \quad (9-8a)$$

¹ Поскольку значение $R_{\text{вых. мин}}$ соответствует большим токам, сопротивление r_6 , входящее в (9-7а), значительно меньше номинального значения вследствие модуляции сопротивления базы (см. § 4-6).

В тех случаях, когда $r_3 \ll R_n \parallel R_3$,

$$K_u = \frac{(1 + \beta) (r_k^* \parallel R_3 \parallel R_n)}{R_r + r_6 + (1 + \beta) (r_k^* \parallel R_3 \parallel R_n)}. \quad (9-86)$$

Из (9-8) ясно видно, что $K_u < 1$ и что величина K_u в первую очередь определяется сопротивлением R_r . При $R_r = 0$ коэффициент передачи обычно составляет не менее 0,995, но уже при $R_r = 100 \div 300$ Ом он падает до 0,98—0,99, а при $R_r = 1 \div 2$ кОм не превышает 0,95—0,98. Коэффициент передачи существенно зависит от сопротивления $R_3 \parallel R_n$. Сопротивление r_k^* обычно не играет большой роли, но оно принципиально ограничивает величину K_u . Так, если не учитывать внутреннюю обратную связь по напряжению, полагая $\mu_{\text{эк}} = 0$, то при $R_r \rightarrow 0$ и $R_3 \parallel R_n \rightarrow \infty$ получаем:

$$K_{u \text{ макс}} \approx \frac{r_k}{r_k + r_6}. \quad (9-9)$$

Это предельное значение доходит до 0,9995 и больше.

Ослабляя входное напряжение, повторитель усиливает ток и мощность. Коэффициент усиления тока нетрудно получить из рис. 9-2, переходя на входе к генератору тока I_r , шунтированному сопротивлением R_r , и считая $r_3 \ll r_k^*$:

$$K_i = \frac{I_n}{I_r} = (1 + \beta) \frac{R_r}{R_r + R_{\text{вх}}} \frac{R_3 \parallel r_k^*}{(R_3 \parallel r_k^*) + R_n}. \quad (9-10)$$

Как видим, усиление тока растет с увеличением сопротивления источника сигнала и уменьшением сопротивления нагрузки. При условиях $R_n \ll R_3 \parallel r_k^*$ и $R_r \gg R_{\text{вх}}$ получается $K_{i \text{ макс}} \approx (1 + \beta)$.

Перейдем к вопросу о передаче импульсов. Из рис. 9-3 видно, что если бы емкости C_k и C_n отсутствовали, то в момент поступления ступенчатого сигнала E_r на выходе сразу появилось бы некоторое напряжение, обусловленное непосредственной передачей сигнала через делитель, состоящий из сопротивлений $(R_r + r_6)$ и $(R_3 \parallel R_n)^*$. Никакого усиления мощности при такой передаче, разумеется, нет. Это подтверждается формулой (9-10), которая в первый момент, при $\beta = 0$, дает $K_i < 1$. Затем, по мере нарастания функции $\beta(t)$, часть тока βI_6 ответвлялась бы на выход и выходное напряжение дополнительно увеличивалось бы вплоть до установившегося значения, определяемого формулами (9-8) (кривая $C_k = 0, C_n = 0$ на рис. 9-4). Другая часть нарастающего тока βI_6 ответвлялась бы в цепь источника сигнала навстречу исходному току $I_6(0)$. Постепенное уменьшение входного тока соответствовало бы росту усиления и мощности вплоть до установившихся значений.

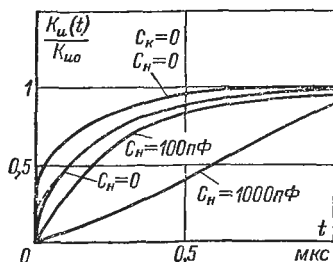


Рис. 9-4. Переходные характеристики повторителя в области малых времен.

* В начале переходного процесса, когда $\beta = 0$, сопротивление $r_k^*(0) = r_k$ и может не учитываться. Сопротивлением r_3 также можно пренебречь.

Влияние емкости C_K^* сказывается в том, что начальный скачок напряжения на выходе отсутствует. Однако если постоянная времени, обусловленная емкостью, значительно меньше, чем τ_β (а это часто бывает так), то начальный участок переходной характеристики будет более крутым, чем последующий, обусловленный диффузионными процессами (кривая $C_H = 0$ на рис. 9-4). Емкость нагрузки C_H в первом приближении оказывает такое же влияние, как и емкость C_K^* , т. е. еще больше уменьшает крутизну начального участка фронта (кривая $C_H = 100$ пФ на рис. 9-4). Кроме того, наличие емкостей приводит к увеличению времени нарастания в целом.

Для количественного анализа передачи импульсов подставим в формулу (9-8б) изображения $\beta(s)$, $Z_K^*(s)$ и $Z_{\beta, H}(s)$. Тогда после преобразований можно привести изображение K_H к следующей приближенной форме:

$$K_H(s) = K_{H0} \frac{1 + s\tau_\alpha}{1 + \frac{s a_1 + s^2 a_2}{1 + \beta_{oe} \gamma_6}}. \quad (9-11)$$

Здесь приняты обозначения:

$$\begin{aligned} a_1 &= \tau'_\beta + \tau'_K + \tau'_{H, \beta}; \\ a_2 &= \tau_\beta \tau'_{H, \beta} + \tau_\alpha \tau'_K; \\ \tau'_\beta &= \tau_\beta \gamma_K^*; \quad \tau'_{H, \beta} = \tau_{H, \beta} \gamma_K^*; \quad \tau'_K = \tau_K (1 - \gamma_K^*). \end{aligned}$$

Коэффициенты токораспределения γ_6 и γ_K^* и коэффициент усиления β_{oe} определяются по тем же принципам, что и ранее (см. гл. 7):

$$\begin{aligned} \gamma_6 &= \frac{R_\beta \parallel R_H}{R_\beta \parallel R_H + (R_\Gamma + r_6)}; \\ \gamma_K^* &\approx \frac{r_K^*}{r_K^* + R_\beta \parallel R_H \parallel (R_\Gamma + r_6)}; \\ \beta_{oe} &= \beta \frac{r_K^*}{r_K^* + R_\beta \parallel R_H \parallel (R_\Gamma + r_6)}. \end{aligned}$$

Легко заметить, что постоянные времени $\tau'_{H, \beta}$ и τ'_K являются произведениями соответствующей емкости (C_H или C_K^*) на сопротивление всех параллельно соединенных ветвей эквивалентной схемы:

$$R_\beta \parallel R_H \parallel r_K^* \parallel (R_\Gamma + r_6).$$

Из выражения (9-11) следует, что переходный процесс может иметь не только апериодический, но и колебательный характер.

Условием апериодического режима [как можно убедиться, решив квадратное характеристическое уравнение для знаменателя (9-11)] является неравенство

$$\frac{1}{2} (\tau'_\beta + \tau'_K + \tau'_{H, \beta}) > \sqrt{(\tau_\beta \tau'_{H, \beta} + \tau_\alpha \tau'_K) (1 + \beta_{oe} \gamma_6)}. \quad (9-12)$$

Это неравенство требует довольно определенного сочетания параметров, в частности либо достаточно малой, либо достаточно большой емкости нагрузки [126]. При «средних» значениях C_H переходная характеристика может иметь выбросы (кривая $C_H = 1000$ пФ на рис. 9-4). Разумеется, использовать большую емкость C_H для ликвидации выбросов было бы неправильно, так как при этом

возрастает длительность фронта. Анализируя выражение (9-11), можно показать, что в случае чисто активной нагрузки ($C_n = 0$) переходный процесс всегда будет аperiodическим.

Выражение для времени нарастания фронта в общем виде получается слишком сложным, хотя в принципе его легко найти из формулы (9-11). Мы ограничимся частным случаем, когда $C_n = 0$. При этом можно пренебречь квадратичным членом в знаменателе (9-11) и получить приближенно:

$$K_u(s) = K_{u0} \frac{1 + s\tau_\alpha}{1 + s \frac{\tau'_\beta + \tau'_\kappa}{1 + \beta_{oe}\gamma_6}}. \quad (9-13)$$

При $t = 0$ (т. е. при $s = \infty$) эта функция имеет конечную величину, что соответствует начальному скачку, точнее — крутому начальному участку переходной характеристики. Из (9-13) по общим правилам (см. сноску на с. 199) находим время нарастания:

$$t_n = 2,2 \sqrt{\left(\frac{\tau'_\beta + \tau'_\kappa}{1 + \beta_{oe}\gamma_6}\right)^2 - \tau_\alpha^2}. \quad (9-14)$$

Если $R_r > R_s \parallel R_n$, т. е. $\gamma_6 < 1/2$, то в выражении (9-14) можно пренебречь слагаемым τ_α^2 . Тогда

$$t_n \approx 2,2 \frac{\tau'_\beta + \tau'_\kappa}{1 + \beta_{oe}\gamma_6}. \quad (9-15a)$$

Если, наоборот, $R_r < R_s \parallel R_n$, т. е. $\gamma_6 \approx 1$, то можно положить $\tau'_\beta / (1 + \beta_{oe}\gamma_6) = \tau_\alpha$. Тогда

$$t_n \approx 2,2 \frac{\sqrt{(2\tau'_\beta + \tau'_\kappa) \tau'_\kappa}}{1 + \beta_{oe}\gamma_6}. \quad (9-15б)$$

Например, если $\tau'_\beta = 0,2$ мкс; $\tau'_\kappa = 0,1$ мкс; $\beta_{oe} = 60$, то при $\gamma_6 = 0,25$ из формулы (9-15a) получим $t_n \approx 45$ нс. При $\gamma_6 = 1$ из формулы (9-15б) получим $t_n \approx 12$ нс. Данный пример, а в общем виде и выражение (9-14) подтверждают, что *время нарастания уменьшается при малых сопротивлениях R_r .*

Частотные характеристики повторителя получаются из формул (9-11) и (9-13) заменой оператора s на $j\omega$. Получающаяся характеристика $K_u(\omega)$ при $C_k = 0$ и $C_n = 0$ не падает до нуля на высоких частотах (рис. 9-5). Этот случай соответствует наличию начального скачка на переходной характеристике. При $C_k \neq 0$ и $C_n = 0$ характеристика $K_u(\omega)$ тоже как бы стремится к некоторому пределу (на рис. 9-5 показан пунктиром), однако на высоких частотах все же падает до нуля. В результате кривая $K_u(\omega)$ имеет излом — «ступеньку», которая соответствует двум участкам с разными постоянными времени на переходной характеристике. При $C_n \neq 0$ кривые $K_u(\omega)$ имеют либо плавный характер, либо [в случае немонотонной переходной характеристики, когда не выполняется условие (9-12)] подъем на сравнительно небольшой частоте (рис. 9-5). Граничная

частота с увеличением емкости непрерывно уменьшается. В частном случае, при $C_H = 0$ и достаточно большом R_T , пренебрегая членом $\sigma\tau_\alpha$ в числителе (9-13), получаем:

$$\omega_B \approx \frac{1 + \beta_{oe}\gamma_6}{\tau'_B + \tau'_K}. \quad (9-16)$$

Например, при $\tau'_B = 0,2$ мкс; $\tau'_K = 0,1$ мкс и $\beta_{oe}\gamma_6 = 20$ получим $f_B \approx 10$ МГц.

Динамический диапазон. Эмиттерный повторитель допускает работу с значительно большими входными сигналами, чем обычный каскад ОЭ. Действительно, поскольку потенциал эмиттера «следит» за потенциалом базы, последний может меняться примерно от нуля до напряжения питания. Более точно этот диапазон определяется следующими соображениями:

1. При напряжении U_6 , близком к нулю, т. е. в области малых токов, сопротивление r_3 согласно (4-22) сильно возрастает. Соответственно уменьшается коэффициент

K_u . Считая, что условием постоянства K_u является $r_3 \ll \ll R_3 \parallel R_H$, можно, используя (4-22), записать:

$$I_{э. мин} = \frac{\varphi_T}{r_{э. макс}} \gg \frac{\varphi_T}{R_3 \parallel R_H}.$$

Определив ток $I_{э. мин}$, трудно по вольт-амперным характеристикам оценить значение $U_{6. мин}$.

2. В области малых коллекторных напряжений, т. е. при U_6 , близком к E_K , уменьшается коэффициент передачи β ,

что приводит к снижению K_u и увеличению времени нарастания фронта. Поэтому значение $U_{6. макс}$ обычно ограничивается условием $|U_{K6}| > 1$ В.

Нелинейность амплитудной характеристики $U_{вых} (U_{вх})$ в диапазоне $U_{6. мин} - U_{6. макс}$ обусловлена согласно (9-8) изменениями параметров β и r_K^* . При небольших значениях R_T эта нелинейность обычно не превышает нескольких процентов.

При достаточно больших импульсах с крутыми фронтами эмиттерный повторитель может неодинаково передавать положительные и отрицательные перепады напряжения [127]. Причина такой асимметрии объясняется следующим образом: большой входной импульс с крутым фронтом при наличии емкости C_H может на некоторое время запереть эмиттерный переход транзистора (рис. 9-6, а). Отпирание транзистора и возвращение его в нормальный режим повторения произойдут лишь тогда, когда емкость C_H достаточно разрядится, причем разряд происходит с довольно значительной

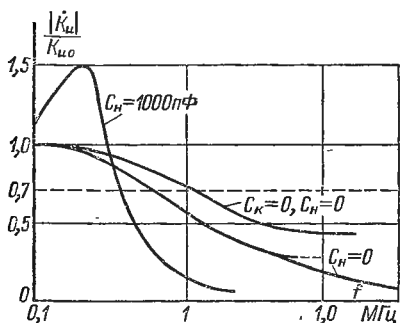


Рис. 9-5. Амплитудно-частотные характеристики повторителя в области высших частот.

постоянной времени $C_H (R_3 \parallel R_H)$. Таким образом, соответствующий фронт выходного импульса будет искажен. В случае транзистора $p-n-p$ искажается положительный фронт (рис. 9-6, б), а в случае транзистора $n-p-n$ — отрицательный.

Описанных искажений не будет, если начальная скорость нарастания входного сигнала будет меньше, чем начальная скорость разряда емкости C_H :

$$\left(\frac{dE_r}{dt}\right)_0 < \frac{U_{30}}{\tau_{\text{ВЫХ}}} \approx \frac{U_{30}}{C_H (R_{\text{ВЫХ}} \parallel R_H)}.$$

Выходное сопротивление $R_{\text{ВЫХ}}$, как уже отмечалось, во время переходного процесса меняется. Поэтому при использовании приведенного условия следует ориентироваться на некоторое промежуточное значение этого сопротивления. Для входных сигналов,

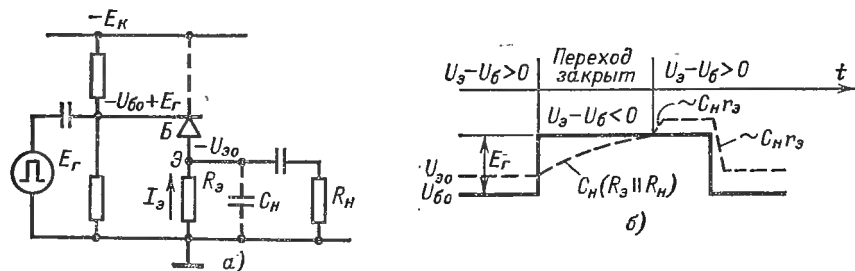


Рис. 9-6. Передача положительного и отрицательного фронтов импульса при наличии нагрузочной емкости.

а — эквивалентная схема; б — временная диаграмма.

составляющих несколько вольт, и емкости нагрузки около 100 пФ типичной допустимой скоростью изменения входного сигнала является 10 В/мкс.

9-3. СЛОЖНЫЕ ПОВТОРИТЕЛИ

Если большое входное сопротивление является первоочередным требованием, то хорошие результаты можно получить с помощью так называемых сложных повторителей, в которых используются два транзистора и более. Следует заметить, что в таких схемах не имеет смысла ставить делитель в цепи базы, так как он не позволяет реализовать то большое входное сопротивление, ради которого применяют сложный повторитель. Поэтому в последующих схемах делитель не показан и считается, что источник сигнала связан с базой повторителя непосредственно, без реактивных элементов. При этом предполагается, что источник сигнала имеет постоянную составляющую напряжения, необходимую для нормального режима повторителя. В противном случае нужно использовать на входе делитель со следящей связью, о котором будет сказано ниже (см. рис. 9-9).

Повторитель на составном транзисторе (по схеме Дарлингтона).

В этом повторителе (рис. 9-7) входное сопротивление 2-го транзистора, вычисляемое по формуле (9-1), играет роль эмиттерной нагрузки по отношению к 1-му транзистору. Следовательно, чтобы найти входное сопротивление повторителя в целом, нужно заменить в формуле (9-16) величину $R_э \parallel R_n$ на $R_{вх2}$; тогда

$$R_{вх} \approx (1 + \beta_1) (r_{к1}^* \parallel R_{вх2}).$$

Если $R_{вх2} > r_{к1}^*$, как обычно бывает, то входное сопротивление повторителя согласно (9-2) будет близко к $r_{к1}$, причем это максимальное значение в отличие от простого повторителя получается при вполне реальном конечном значении $R_э \parallel R_n$ (сотни ом и более).

Количественный расчет параметров $R_{вх}$, K и $R_{вых}$ можно осуществлять либо «покасадно» (рассматривая транзистор T_2 как нагрузку для T_1 , а транзистор T_1 как источник сигнала для T_2), либо заменяя комбинацию T_1, T_2 одним составным транзистором со свойствами ему параметрами (см. § 4-10).

Рис. 9-7. Повторитель на составном транзисторе.

В частности, если использовать второй путь, то, умножая обе части (4-118б) на $1 + \beta_2 \approx \beta_1 \beta_2$ и учитывая соотношение $r_{к1}^*/r_{к2}^* = 1 + \beta_2$ [поскольку $I_{э2} = (1 + \beta_2) I_{э1}$ (см. с. 234)], получаем:

$$R_{вх, макс} = r_{к\Sigma} \approx \frac{1}{2} r_{к1}.$$

Коэффициент передачи в составном повторителе выражается формулой, близкой к (9-9), т. е. он намного ближе к единице, чем в простом повторителе. Значения $K \geq 0,995$ можно считать типичными. Выходное сопротивление, напротив, оказывается значительно меньше, чем в простом повторителе, поскольку истинное значение R_r «пересчитывается» 1-м транзистором согласно (9-6б), так что для 2-го транзистора источник сигнала имеет сопротивление, близкое к $r_{э1}$. Учитывая различие в токах транзисторов, нетрудно прийти к выводу, что результирующее выходное сопротивление составляет примерно $2r_{э2}$. Если же токи обоих транзисторов одинаковы¹, то результирующее выходное сопротивление определяется формулой (9-6в), т. е. $r_{э2} \approx r_{э1}$.

Переходные и частотные характеристики входного сопротивления определяются в первую очередь постоянной времени τ_{oe}

¹ Равенство токов, как отмечалось в § 4-10, достигается путем включения в схему генератора тока (на рис. 9-7 показан пунктиром). Этот генератор, рассчитываемый из условия желательной разности токов $I_{э1} - I_{с2}$, может быть реализован либо в виде резистора, питаемого от источника положительной э. д. с., либо в виде транзистора, включенного по схеме ОБ с заданным током эмиттера.

наиболее низкочастотного из двух транзисторов (им обычно оказывается транзистор T_2 , поскольку он работает с большими токами). Поэтому даже при использовании дрейфовых транзисторов граничная частота (на уровне $0,7 R_{вх}$) редко превышает 500—600 кГц.

Наконец, нужно заметить, что входное сопротивление может сильно меняться с изменением температуры. Это объясняется в первую очередь зависимостями $r_k(T)$ и $\beta(T)$.

В случае германиевых транзисторов добавляется зависимость $I_{с1}(T)$: при повышенной температуре ток $I_{с2}$ заметно уменьшается из-за роста теплового тока, а это означает уменьшение тока $I_{с1}$ и соответствующие изменения параметров $r_{с1}$ и β_1 . В зависимости от выбора начальной рабочей точки транзистора T_1 уменьшение тока $I_{с1}$ может привести и к уменьшению, и к увеличению $R_{вх}$.

Составной повторитель с внутренней обратной связью. Для того чтобы искусственно увеличить сопротивление $r_{к1}$ в схеме на рис. 9-7, нужно изменить потенциал $U_{к1}$ пропорционально потенциалу $U_{с1}$. Тогда разность потенциалов $U_{кс1}$ будет значительно меньше, чем $U_{с1}$, ток через сопротивление $r_{к1}$ сильно уменьшится, а это равносильно увеличению $r_{к1}$. Для реализации этой идеи в схему (рис. 9-8) вводится сопротивление $R_{к1}$ и на коллектор транзистора T_1 полностью подается переменная составляющая выходного напряжения [128]. Батарея E компенсирует постоянную составляющую коллекторного потенциала $U_{к1}$. Практически вместо батареи используют либо конденсатор большой емкости, либо полупроводниковый стабилитрон; однако для расчетов это не имеет значения.

Очевидно, что эквивалентное увеличение сопротивления $r_{к1}$ обусловлено уменьшением тока через него при одном и том же входном сигнале. Действительно, ток $I_{r_{к1}}$ в схеме на рис. 9-8 равен:

$$(E_{г} - U_{вых})/r_{к1} = E_{г}(1 - K_u)/r_{к1}.$$

Поэтому, определяя эквивалентное сопротивление $r_{к1 экв}$ как $E_{г}/I_{r_{к1}}$, получаем:

$$r_{к1 экв} = \frac{r_{к1}}{1 - K_u}. \quad (9-17)$$

Поскольку схема является повторителем (к тому же составным), коэффициент передачи K_u близок к единице и сопротивление $r_{к1}$ увеличивается в сотни раз. Положим для простоты $r_{к1} = \infty$. Тогда из (4-1186) получаем $r_{к\Sigma}^* = r_{к2}^*$ и, следовательно,

$$r_{к\Sigma} \approx \beta_{\Sigma} r_{к\Sigma}^* \approx \beta_1 r_{к2}. \quad (9-18)$$

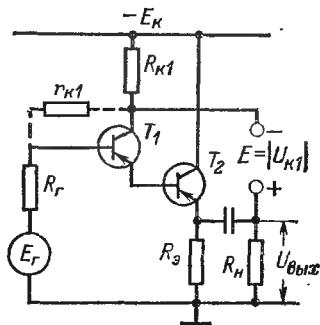


Рис. 9-8. Составной повторитель с внутренней обратной связью.

Таким образом, в данной схеме можно получить значение $R_{\text{вх. макс}}$ значительно большее, чем в простом повторителе. Для этого, однако, нужно, чтобы рабочий ток транзистора T_2 был таким же, как в простом повторителе (тогда будут одинаковы коллекторные сопротивления). Практически входное сопротивление в схеме на рис. 9-8 на низких частотах может достигать 100 МОм, если, конечно, сопротивления $R_3 \parallel R_n$ и $R_{\text{к1}}$ достаточно велики.

С ростом частоты (или скорости изменения сигнала) входное сопротивление уменьшается. Причина этого уменьшения заключается не только в зависимости $\beta(f)$, но и в запаздывании сигнала обратной связи. В самом деле, в первый момент $U_{\text{вых}}(0) = 0$,

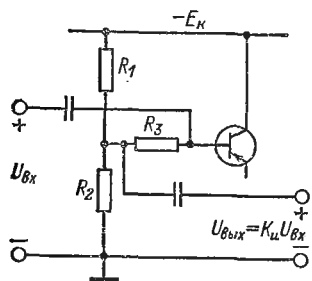


Рис. 9-9. Вход повторителя со следящей связью.

поэтому сопротивление $r_{\text{к1}}$ имеет нормальное, сравнительно небольшое значение. Установившаяся величина (9-17) получится лишь после установления коэффициента K_u . Под действием обоих инерционных факторов $\beta(f)$ и $K_u(f)$ входное сопротивление в данном повторителе устанавливается медленнее, чем в предыдущем, а его граничная частота оказывается соответственно меньше.

В связи с огромным входным сопротивлением сложных повторителей особенно остро встает вопрос о цепи смещения базы: делать сопротивления делителя R_1, R_2 (см. рис. 9-1) порядка десятков мегаом нельзя не только из-за температурной нестабильности, но главным образом из-за невозможности обеспечить нужный ток базы. В случае усилителей переменного тока наилучшим выходом из положения является использование *следящей связи* в цепи базы (рис. 9-9). Идея такого решения та же, что и в схеме на рис. 9-8: сопротивление R_3 выбирается сравнительно низкоомным (сотни килоом), способным обеспечить необходимый ток базы, а по отношению к переменным составляющим величина R_3 искусственно повышается в $1/(1 - K_u)$ раз за счет обратной связи с выхода повторителя. Таким образом, эквивалентное сопротивление $R_{\text{зэкв}}$ может достигать десятков мегаом и не будет существенно шунтировать вход повторителя. Делитель R_1, R_2 в такой схеме может иметь любое разумное сопротивление.

Повторитель с динамической нагрузкой. Как в простом, так и в составном повторителе увеличение сопротивления R_3 затрудняется ростом постоянной составляющей напряжения $I_3 R_3$. Это затруднение можно обойти, используя вместо омического сопротивления R_3 нелинейный элемент с большим дифференциальным сопротивлением и малым сопротивлением постоянному току. Для этого можно использовать второй транзистор, включенный по схеме ОБ или ОЭ. Последнее включение (рис. 9-10) позволяет

легко осуществить смещение транзистора T_2 от того же источника питания E_K .

Из рис. 9-10 видно, что коллекторные токи обоих транзисторов одинаковы. Следовательно, сопротивление R_6 можно выбрать из условия $E_K/R_6 \approx I_{62} \approx I_{K1}/\beta_2$, где I_{K1} — желательный ток основного транзистора. Напряжение U_{K32} близко к потенциалу базы U_{61} , задаваемому тем или иным известным способом, и может быть достаточно малым.

Дифференциальные параметры схемы описываются теми же формулами, что и для простого повторителя, если заменить в них R_6 на r_{K2}^* . При этом, конечно, должно быть выполнено условие $R_H > > r_{K2}^*$, без которого данный повторитель не дает желаемого эффекта. Положим для простоты $R_H = \infty$ и будем считать параметры обоих транзисторов одинаковыми. Тогда из (9-1) получим:

$$R_{вх} \approx \frac{r_K}{2}. \quad (9-19)$$

Отсюда видно, что входное сопротивление имеет весьма значительную величину и не зависит от коэффициента β . Поэтому, в частности, характеристики $z_{вх}(t)$ и $Z_{вх}(\omega)$ определяются практически только постоянной времени τ_K . В этом легко убедиться, заменив в формуле (9-19) сопротивление r_K на $Z_K = r_K/(1 + s\tau_K)$. Кроме того, на величине $R_{вх}$ не будет сказываться температурная зависимость $\beta(T)$, что особенно ценно для кремниевых транзисторов.

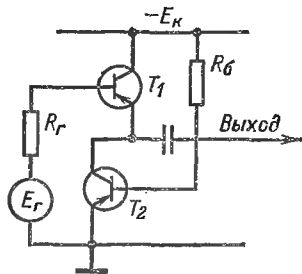


Рис. 9-10. Повторитель с «динамической» нагрузкой.

Глава десятая

КАСКАД С ЭМИТТЕРНЫМ ВХОДОМ

Как уже отмечалось, каскад с эмиттерным входом, т. е. каскад ОБ, имеет более ограниченное применение, чем каскады ОЭ и ОК. Познакомившись с параметрами каскада ОБ (рис. 10-1), мы уточним эти ограничения и их причины.

10-1. ОБЛАСТЬ СРЕДНИХ ЧАСТОТ

На рис. 10-2 показана упрощенная эквивалентная схема каскада в области средних частот. На схеме опущены малосущественные сопротивления R_6 и r_K . Последнее в десятки раз превышает значение R_K и поэтому учет его обычно не оправдан.

Структурное подобие эквивалентных схем каскадов ОБ и ОЭ (см. рис. 7-2) позволяет воспользоваться простой заменой параметров в соответствующих выражениях, выведенных для каскада ОЭ

в § 7-2. А именно, достаточно заменить параметры β , r_3 , r_6 , r_k^* соответственно на $-\alpha$, r_6 , r_3 , r_k .

Произведя такую замену в формулах (7-2), (7-4а) и (7-7а), получим:

$$R_{\text{вх}} = r_3 + (1 - \alpha) r_6; \quad (10-1)$$

$$K_u = \alpha \frac{R_k \parallel R_n}{R_r + R_{\text{вх}}}; \quad (10-2)$$

$$K_i = \alpha \frac{R_r}{R_r + R_{\text{вх}}} \frac{R_k}{R_k + R_n}. \quad (10-3)$$

Отсюда видно, что входное сопротивление у каскада ОБ в $(1 + \beta)$ раз меньше, чем у каскада ОЭ (не больше 30 Ом при токе 1 мА), а усиление по току меньше единицы. Что касается усиления

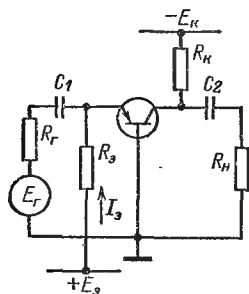


Рис. 10-1. Принципиальная схема каскада ОБ.

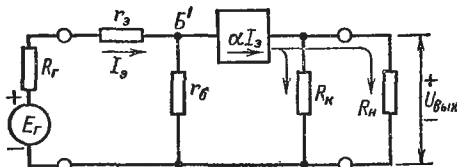


Рис. 10-2. Эквивалентная схема каскада ОБ в области средних частот.

по напряжению, то оно зависит от сопротивления нагрузки и в общем случае может быть большим. Однако в частном, но очень важном случае, когда нагрузкой служит вход такого же каскада, коэффициент усиления по напряжению оказывается меньше единицы. Это ясно видно из формулы (10-2), если положить $R_n = R_{\text{вх}}$. Поэтому при работе одного каскада ОБ на другой не получается ни усиления тока, ни усиления напряжения и, следовательно, такое соединение не имеет смысла.

Из сказанного следует, что каскад ОБ может быть усилителем только в следующих трех случаях: при использовании трансформаторной связи¹, при работе на каскад с более высокоомным входом (ОЭ или ОК) и в качестве выходного каскада, работающего на достаточно большое сопротивление нагрузки. Первая из перечисленных

¹ Как известно, трансформатор позволяет «преобразовывать» сопротивления (см. гл. II). В частности, он может искусственно повысить входное сопротивление следующего каскада, что необходимо для получения $K_u > 1$ в усилителе ОБ.

возможности реализовать в мощных (см. гл. 12) и резонансных усилителях¹.

Еще одна существенная особенность каскада ОБ состоит в том, что он не вносит сдвига фаз между выходными и входными величинами (токами и напряжениями); это следует из положительных знаков коэффициентов K_u и K_i .

Как и в каскаде ОЭ, можно представить параметры схемы в форме, отражающей внутреннюю обратную связь. Произведя замену параметров в формулах (7-15) и (7-16) (см. начало с. 396), получим:

$$R_{вх} = (r_s + r_6) (1 - \alpha \gamma_{s0}); \quad (10-4)$$

$$K_u = \frac{\alpha}{1 - \alpha \gamma_s} \frac{R_k \parallel R_H}{R_r + r_s + r_6}. \quad (10-5)$$

Здесь $\gamma_s = \frac{r_6}{R_r + r_6 + r_s}$ — коэффициент обратной связи по току, аналогичный коэффициенту γ_6 в схеме ОЭ*.

Из выражений (10-4) и (10-5) видно, что внутренняя обратная связь в каскаде ОБ положительна, так как она уменьшает входное сопротивление и увеличивает усиление по напряжению по сравнению со случаем отсутствия обратной связи (когда левый зажим генератора αI_s отключен от точки B' и заземлен).

Выходное сопротивление каскада согласно рис. 10-2 равно R_k . Если учесть коллекторное сопротивление r_k , то, произведя замену параметров в формуле (7-22), получим:

$$R_{вых} = [r_k (1 - \alpha \gamma_s)] \parallel R_k. \quad (10-6)$$

Сравнивая формулы (10-5) и (10-6) с соответствующими формулами для каскада ОЭ, легко убедиться, что они совпадают в одном частном случае, когда сопротивление источника сигнала равно нулю. Это объясняется тем, что при $R_r = 0$ входные цепи обоих каскадов оказываются идентичными, т. е. различие между понятиями «общая база» и «общий эмиттер» стирается. Сохраняются, однако, особенности каскада ОБ — малое входное сопротивление и отсутствие сдвига фаз между входным и выходным напряжениями.

В тех случаях, когда нужно учесть влияние сопротивления r_k , можно воспользоваться методикой, изложенной в § 7-2 для каскада ОЭ, и заменить в соответствующих формулах коэффициент α на $\alpha_{oe} = \alpha \gamma_k$, где $\gamma_k = r_k / (r_k + R_k \parallel R_H)$.

Для маломощных транзисторов, у которых $r_k \approx 1$ МОм, такая поправка редко необходима.

В области низших частот различие между каскадами ОБ и ОЭ чисто количественное и не очень существенное. Формулы, выведенные в § 7-3, остаются в силе и не нуждаются в специальном анализе.

¹ Резонансные усилители в данном издании книги (как и в предыдущих) опущены. Подробный анализ их проведен в работах [67, 69, 129—131].

* Заметим, что γ_s в схеме ОБ и γ_6 в схеме ОЭ не связаны соотношением $\gamma_s = 1 - \gamma_6$, так как сопротивление R_r в первом случае включено в эмиттерную цепь, а во втором — в базовую.

Рассмотрим качественно переходный процесс в схеме на рис. 10-3, пренебрегая сначала емкостью C_K . Предположим, что на входе задана ступенька тока I_3 . Генератор αI_3 в первый момент бездействует (инжектированные носители не дошли до коллектора), и выходные ток и напряжение равны нулю; входное сопротивление в этот момент равно $r_3 + r_6$. В дальнейшем по мере роста $\alpha(t)$ входное сопротивление уменьшается, а выходные ток и напряжение возрастают до своих установившихся значений, соответствующих средним частотам. Скорость этих изменений определяется постоянной времени τ_α , с которой увеличивается ток αI_3 .

Если сопротивление источника сигнала имеет конечную величину R_r , то часть тока αI_3 будет ответвляться в это сопротивление. В данном случае входной ток нарастает вместе с выходным (обратная связь положительна) и, следовательно, переходные процессы затягиваются. В предельном случае, при $R_r = 0$ (генератор э. д. с. на входе), длительность и характер переходных процессов не должны отличаться от длительности и характера их в каскаде ОЭ, так как входные цепи обеих схем при этом делаются идентичными.

Таким образом, с точки зрения переходных и частотных свойств вариант $R_r = 0$, наилучший для каскада ОЭ, будет наилучшим для каскада ОБ (рис. 10-4).

Влияние емкости C_K проявляется в следующем. Во-первых, в момент подачи сигнала часть его через емкость C_K поступает непосредственно на выход. Во-вторых, последующий переходный процесс определяется не только постоянной времени τ_α , но и постоянной времени емкости. Поэтому переходные и частотные свойства ухудшаются.

Для количественного анализа воспользуемся снова структурной аналогией между эквивалентными схемами каскадов ОБ и ОЭ и заменим в формуле (7-39) величины β , r_3 и др. на $-\alpha$, r_6 и т. д., а также τ_β на τ_α и C_K^* на C_K (последняя замена не совсем правомерна, так как C_K^* — операторная, а C_K — действительная величина). Тогда приближенные изображения параметров K_u и K_i получатся в следующей общей форме:

$$A(s) = \frac{A_0}{1 + \frac{s\tau_{\alpha oe}}{1 - \alpha\gamma_3}}, \quad (10-7)$$

где A_0 — значение параметров на средних частотах; $\tau_{\alpha oe} = \tau'_\alpha + \tau'_k$ — эквивалентная постоянная времени;

$$\tau'_\alpha = \tau_\alpha \gamma_k \approx \tau_\alpha; \quad \tau'_k = \tau_k (1 - \gamma_k) \approx C_k (R_k \parallel R_n).$$

Время нарастания согласно общему определению [62] равно:

$$t_n = 2,2 \frac{\tau'_\alpha + \tau'_k}{1 - \alpha\gamma_3}. \quad (10-8)$$

Из формулы (10-7) получают упрощенные частотные характеристики параметров:

$$\dot{A} = \frac{A_0}{1 + j \frac{\omega}{\omega_B}}, \quad (10-9)$$

где верхняя граничная частота

$$\omega_b = \frac{1 - \alpha \gamma_{\beta 0}}{\tau'_\alpha + \tau'_k} \quad (10-10)$$

Сравнение полученных выражений с соответствующими выражениями для каскада ОЭ подтверждает преимущества каскада ОБ в отношении переходных и частотных свойств. Например, при $f_\alpha = 50$ МГц; $r_b = 25$ Ом; $r_6 = 100$ Ом; $C_k = 2$ пФ; $R_k \parallel R_H = 2$ кОм; $R_r = 100$ Ом в каскаде ОБ получаем $t_H \approx 35$ нс; $f_b = 10$ МГц, тогда как для схемы ОЭ при $\beta = 50$ будем иметь $t_H \approx 120$ нс; $f_b \approx 2,8$ МГц.

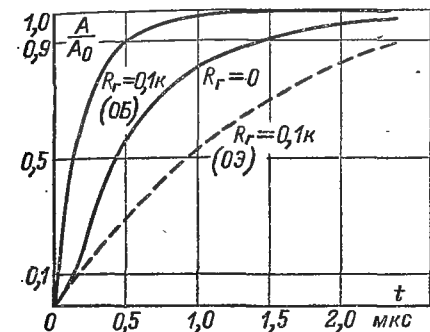


Рис. 10-4. Сравнительные переходные характеристики каскадов ОБ и ОЭ при $R_r = 0$ и $R_r \neq 0$.

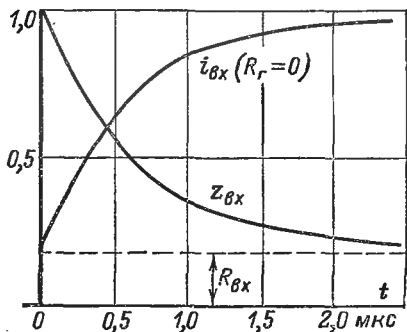


Рис. 10-5. Переходные характеристики входного тока и входного сопротивления в каскаде ОБ.

Упрощенное изображение входного сопротивления получается из формулы (10-4) при подстановке $\alpha(s)$. После некоторых преобразований оно может быть представлено в следующей форме:

$$Z_{вх}(s) \approx R_{вх} \frac{1 + s \frac{\tau_\alpha}{1 - \alpha \gamma_{\beta 0}}}{1 + s \tau_\alpha} \quad (10-11)$$

Нетрудно убедиться, что при $t = 0$ (т. е. при $s = \infty$) начальное значение входного сопротивления больше, чем $R_{вх}$. Соответственно входной ток (при ступенчатом сигнале E_r) после начального скачка плавно нарастает до установившегося значения. На рис. 10-5 показаны функции $z_{вх}(t)$ и $i_{вх}(t)$ в относительном масштабе. Из сказанного следует, что *входное сопротивление каскада ОБ имеет индуктивный характер* в отличие от каскадов ОЭ и ОК (т. е. каскадов с базовым входом), у которых оно носит емкостный характер. Естественно, что при синусоидальном сигнале входное сопротивление каскада ОБ увеличивается с ростом частоты.

10-3. КАСКАД С ЭМИТТЕРНОЙ СВЯЗЬЮ

Органическим сочетанием каскадов ОК и ОБ является каскад с эмиттерной связью (рис. 10-6), предназначенный для с и н ф а з н о г о усиления (оба каскада ОК и ОБ не изменяют полярности сигнала).

Анализ каскада проведем только в области средних частот. Для этого воспользуемся формулами, выведенными в § 9-2 и 10-1, считая параметры обоих транзисторов одинаковыми. Кроме того, примем для обоих транзисторов $r_k^* = \infty$ (что упрощает анализ) и положим $R_3 = \infty$. Последнее предположение основано на том, что эмиттерные токи I_{s1} и I_{s2} под действием сигнала E_r меняются в противоположных направлениях, а ток через сопротивление R_3 остается при этом практически неизменным¹.

Нагрузкой повторителя (транзистор T_1) является входное сопротивление каскада ОБ (транзистор T_2). Поэтому, перейдя в формуле (10-1) от α к β и добавив к r_6 сопротивление $R_{62} = R_1' \parallel R_2'$, получим:

$$R_{u1} = R_{u2} = r_3 + \frac{R_{62} + r_6}{1 + \beta}.$$

Теперь из формулы (9-8а), полагая $R_n = R_{u1}$, $R_r \ll R_1' \parallel R_2'$ и опуская R_3 и r_k^* , находим коэффициент передачи повторителя:

$$K_u' = \frac{U_3}{E_r} = \frac{R_{62} + r_6 + (1 + \beta) r_3}{R_r + R_{62} + 2[r_6 + (1 + \beta) r_3]} \quad (10-12a)$$

Рис. 10-6. Каскад с эмиттерной связью.

Коэффициент усиления каскада ОБ найдем из формулы (10-2), полагая в ней $R_r = 0$ (поскольку в данном случае коэффициент усиления определяется по отношению к напряжению непосредственно на эмиттере) и подставляя в нее полученное выше значение R_{u2} :

$$K_u'' = \frac{U_{\text{вых}}}{U_3} = \frac{\alpha (R_k \parallel R_n)}{r_3 + (1 - \alpha) (R_{62} + r_6)} = \frac{\beta (R_k \parallel R_n)}{R_{62} + r_6 + (1 + \beta) r_3} \quad (10-12б)$$

Перемножив коэффициенты K_u' и K_u'' , найдем коэффициент усиления каскада в целом:

$$K_u = \frac{U_{\text{вых}}}{E_r} = \frac{\beta (R_k \parallel R_n)}{R_r + R_{62} + 2[r_6 + (1 + \beta) r_3]} \quad (10-13)$$

Рассмотренная схема обладает сравнительно узким динамическим диапазоном входных сигналов. В самом деле, пусть, например, потенциал U_{62} фиксирован и пусть U_{360} — напряжение на эмиттерных переходах в отсутствие сигналов. Тогда потенциал U_{61} не может увеличиваться (по модулю) больше чем на $2U_{360}$ (иначе запрется транзистор T_2) и уменьшаться больше чем на $2U_{60}$

¹ Подробнее о роли сопротивления R_3 см. в гл. 14.

(иначе заперется транзистор T_1) [при этом считается, что коэффициент передачи повторителя близок к $1/2$, см. (10-12а)]. Таким образом получаем ограничение на входной сигнал: $|E_r| < 2U_{эб0}$. Практически, с учетом нелинейных искажений, $E_{г, макс} \approx (0,1 - 0,2)$ В для германиевых и $(0,4 - 0,5)$ В для кремниевых транзисторов.

По тем же причинам потенциалы баз в режиме покоя необходимо выравнивать, в противном случае один из транзисторов окажется запертым. Отмеченные затруднения заставляют иногда использовать в цепи эмиттеров дополнительные сопротивления — «связки» (см. рис. 14-2).

10-4. КАСКОД

Каскодом называют последовательное соединение каскадов ОЭ и ОБ (рис. 10-7). Коллекторной нагрузкой транзистора T_1 , включенного по схеме ОЭ, является эмиттерная цепь транзистора T_2 , включенного по схеме ОБ.

В отличие от простейшего варианта схемы ОБ (рис. 10-1) база транзистора T_2 не заземлена, а находится под постоянным потенциалом $|E_6| < |E_K|$. Это не меняет свойств каскада ОБ, так как потенциал базы можно условно считать нулевым. Практически потенциал $-E_6$ получают либо от особого источника смещения (если он имеется), либо от делителя напряжения, подключенного между шинкой $-E_K$ и землей. В последнем случае одно из плеч делителя должно быть обязательно зашунтировано достаточно большой емкостью, чтобы база транзистора T_2 могла считаться «заземленной по переменному току» на любой частоте сигнала (или при любых рабочих длительностях импульсов).

Поскольку входное сопротивление каскада ОБ ничтожно мало [см. (10-4)], можно считать, что каскад ОЭ работает в режиме короткого замыкания на выходе. Тогда его усилительная способность характеризуется крутизной (8-17а). При этом коэффициент β_{oe} следует заменить на β , так как не нужно вносить поправку на нагрузку. Считая, кроме того, $\alpha_2 = 1$, приходим к выводу, что результирующая крутизна каскода такая же, как у каскада ОЭ.

Умножая крутизну (8-17а) на суммарную нагрузку $R_K \parallel R_H$, получаем коэффициент усиления напряжения (см. табл. 8-2):

$$K_u = - \frac{\beta_1}{1 + \beta_1 \gamma_{61}} \frac{R_K \parallel R_H}{R_r + r_{61} + r_{э1}}. \quad (10-14)$$

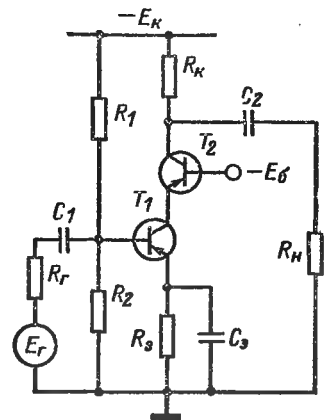


Рис. 10-7. Принципиальная схема каскода.

Выражение (10-14) полностью совпадает с выражением (7-16), которое было выведено без учета сопротивления r_{κ}^* . Значит, можно сказать, что каскод обеспечивает такое же усиление, как и деализированный каскод ОЭ, в котором транзистор имеет бесконечно большое коллекторное сопротивление¹. Иначе говоря, коэффициент усиления у каскода при прочих равных условиях несколько выше, чем у реального каскода ОЭ, поскольку $\beta > \beta_{oe}$. Однако это преимущество каскода не очень существенно по сравнению с другими (см. ниже).

Главная особенность каскода состоит в том, что его выход благодаря наличию каскада ОБ изолирован (или, как говорят, «развязан») от коллектора усилительного транзистора T_1 . В результате тракт прямой передачи сигнала от входа к выходу становится однопроводным или во всяком случае более однонаправленным, чем в обычном каскаде ОЭ.

В частности, «развязка» коллектора T_1 от суммарной нагрузки исключает влияние последней на коэффициент β_1 , о чем говорилось выше.

Другим, не менее важным следствием «развязки» является существенное уменьшение входной емкости, а значит, увеличение входного сопротивления на высоких частотах. Действительно, поскольку при каскодном включении можно, как мы убедились, считать $\gamma_{\kappa 1}^* = 1$, то из формулы (7-36) следует, что в эквивалентной постоянной времени каскада ОЭ исчезает составляющая τ'_{κ} , связанная с коллекторным импедансом. Соответственно входная емкость, определяемая постоянной времени τ_{oe} (рис. 7-11), уменьшится.

Наконец, повышенная однонаправленность в каскоде означает резкое ослабление обратной связи с выхода на вход. Это способствует устойчивой работе схемы, особенно в резонансных (контурных) усилителях, в которых при использовании схемы ОЭ нередко возникает паразитная генерация (самовозбуждение).

Как известно, в ламповой технике аналогичные преимущества обеспечивают лампы с экранирующими сетками по сравнению с триодами. Поэтому каскод как целое можно считать транзисторным аналогом тетродов и пентодов.

Расчет каскода по постоянному току осуществляется раздельно для каскада ОЭ (при напряжении питания — E_6) и для каскада ОБ (при напряжении коллекторного питания — $E_{\kappa} + E_6$ и эмиттерном токе $I_{э2} = I_{\kappa 1}$). Если потенциал — E_6 задается делителем напряжения с сопротивлениями плеч R_3 и R_4 , то в формулах (6-4) и последующих следует полагать:

$$E_6 = E_{\kappa} \frac{R_4}{R_3 + R_4}; \quad R_6 = R_3 \parallel R_4.$$

¹ При выводе (10-14) мы пренебрегли коэффициентом γ_{κ} (см. с. 397), поскольку этот коэффициент заметно отличается от единицы только при сопротивлениях $R_{\kappa} \parallel R_H$ порядка 100 кОм и более, которые в обычных схемах неприемлемы.

Что касается эмиттерной цепи, то ее параметрами будут:

$$E_э = I_{к1} (R_{вых})_{ОЭ}; \quad R_э = (R_{вых})_{ОЭ},$$

где $(R_{вых})_{ОЭ}$ рассчитывается по формуле (7-23).

Расчет каскода на средних и низших частотах неспецифичен и тоже выполняется раздельно для каскадов ОЭ и ОБ. Единственной особенностью является расчет блокирующей емкости в цепи делителя (см. с. 401). Емкость $C_{бл}$ рассчитывается, исходя из желательного значения низкочастотной постоянной времени τ_n (см. § 7-3), которая в данном случае имеет вид¹:

$$\tau_n \approx C_{бл} (R_3 \parallel R_4). \quad (10-15)$$

В области высших частот для анализа можно использовать формулу (7-60), заменив параметры 2-го каскада на параметры схемы ОБ, как это было сделано в § 10-2. Кроме того, оператор s нужно заменить на $j\omega$. Анализ показывает, что частотные характеристики каскода в целом определяются каскадом ОЭ, как более инерционным. Следовательно, эти характеристики можно рассчитывать по формулам § 7-4, имея в виду отмеченную выше особенность ($\gamma_k^* = 1$), связанную с режимом короткого замыкания в коллекторной цепи транзистора T_1 .

Более подробный анализ каскодов можно найти в работе [132], где, однако, используется система y -параметров.

Глава одиннадцатая

УСИЛИТЕЛИ С ТРАНСФОРМАТОРНОЙ СВЯЗЬЮ

11-1. ВВЕДЕНИЕ

На рис. 11-1 показаны две основные схемы каскадов с трансформаторным входом и выходом. Чаще применяется схема на рис. 11-1, а, которую называют схемой с п о с л е д о в а т е л ь н ы м включением выходного трансформатора Tr_2 . Схема с п а р а л л е л ь н ы м включением выходного трансформатора (рис. 11-1, б) нуждается в переходном конденсаторе². В основе последней схемы лежит обычный каскад ОЭ, к которому просто добавлен трансформатор. Каскад с последовательным включением трансформатора более существенно отличается от каскада с емкостными связями:

1) коллектор транзистора непосредственно присоединен к шине E_k по постоянному току (так как активное сопротивление первичной обмотки Tr_2 мало);

¹ Строго говоря, нужно было бы учесть еще и входное сопротивление транзистора T_2 со стороны базы; однако расчет по формуле (8-19) показывает, что оно обычно пренебрежимо велико.

² Схема включения выходного трансформатора, показанная на рис. 11-1, б, называется еще *схемой параллельного питания*.

2) выходное сопротивление каскада в связи с отсутствием R_k равно $r_k^* (1 + \beta \gamma_6)$ [см. (2-23)] и обычно гораздо больше выходного сопротивления в усилителе с параллельным включением, где $R_{вых} \approx R_k$.

Кроме того, при последовательном включении через первичную обмотку трансформатора Tr_2 протекает постоянный подмагничивающий ток.

Что касается входного трансформатора Tr_1 , то способы включения его вторичной обмотки, показанные на рис. 11-1, в сущности равноценны: конденсаторы C_1 и C_3 имеют достаточно большую емкость, так что по переменному току вторичная обмотка присоединена между базой и эмиттером. В обоих случаях параллельно

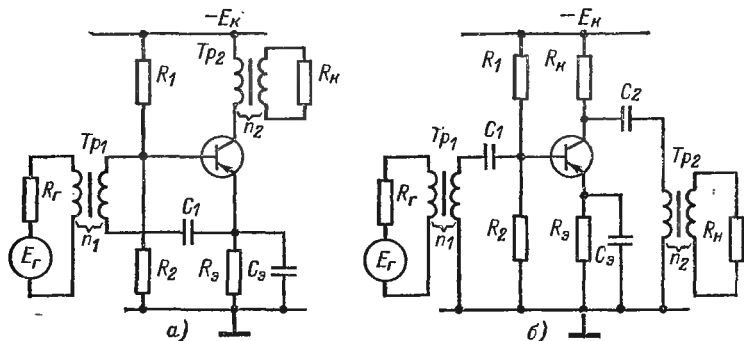


Рис. 11-1. Трансформаторные каскады с последовательным (а) и параллельным (б) включением трансформатора Tr_2 .

входному сопротивлению оказывается включенным эквивалентное сопротивление базового делителя. Если это нежелательно, можно включить обмотку между средней точкой делителя и базой, а одно из плеч делителя зашунтировать большой емкостью.

Применение трансформаторов для связи транзисторных каскадов имеет некоторую специфику по сравнению с ламповыми схемами. Это связано со сравнительно малыми входными сопротивлениями транзисторных каскадов и обратным (по сравнению с лампами) соотношением входного и выходного сопротивлений: $R_{вх} < R_{вых}$. В связи с этим, во-первых, необходимы меньшие индуктивности обмоток при одной и той же нижней граничной частоте (так как постоянная времени L/R получается больше) и, во-вторых, для межкаскадной связи следует обычно применять не повышающие, а понижающие трансформаторы [62]. Последний вывод заслуживает специального рассмотрения.

11-2. КОЭФФИЦИЕНТ ТРАНСФОРМАЦИИ

Максимальная передача мощности от источника сигнала в нагрузку происходит при равенстве их сопротивлений. Если сопротивления $R_г$ и $R_н$ имеют разные значения, их можно согласовать с помощью трансформатора. Такое согласование основано на том,

что в случае идеального трансформатора (без потерь) имеют место соотношения

$$R'_H = \frac{R_H}{n^2}; \quad R'_r = R_r n^2, \quad (11-1)$$

где $n = \omega_2/\omega_1$ — коэффициент трансформации (ω_2, ω_1 — числа витков вторичной и первичной обмоток); R'_H, R'_r — соответственно сопротивление нагрузки, «пересчитанное» в первичную обмотку, и сопротивление источника сигнала, «пересчитанное» во вторичную обмотку.

При согласовании имеем $R'_H = R_r$ или $R'_r = R_H$, что приводит к одному и тому же выражению для согласующего коэффициента трансформации:

$$n_{\text{согл}} = \sqrt{\frac{R_H}{R_r}}. \quad (11-2)$$

Если m -й каскад усилителя является промежуточным, то для выходного трансформатора в полученной формуле нужно использовать величины $R_r = R_{\text{вых}(m-1)}$ и $R_H = R_{\text{вх } m}$, а для выходного трансформатора — величины $R_r = R_{\text{вых } m}$ и $R_H = R_{\text{rx}(m+1)}$. В случае идентичных каскадов коэффициенты трансформации для обоих трансформаторов получаются одинаковыми и равными:

$$n_{\text{согл}} = \sqrt{\frac{R_{\text{вх}}}{R_{\text{вых}}}}. \quad (11-3)$$

Это выражение подтверждает вывод о том, что в транзисторных усилителях, у которых $R_{\text{вх}} < R_{\text{вых}}$, коэффициент трансформации должен быть меньше единицы¹.

В каскаде с параллельным включением трансформатора сопротивление R_k обычно не превышает нескольких килоом. Поэтому можно считать $R_{\text{вых}} \approx R_k$ и $\beta_{oe} \approx \beta$. Тогда коэффициент трансформации рассчитывается по формуле

$$n_{\text{согл}} \approx \sqrt{\frac{R_{\text{вх}}}{R_k}} = \sqrt{\frac{r_6 + (1 + \beta)r_3}{R_k}} \quad (11-4a)$$

и для маломощных каскадов, как правило, лежит в пределах 0,25—0,5.

В каскаде с последовательным включением трансформатора сопротивления $R_{\text{вх}}$ и $R_{\text{вых}}$ оказываются взаимосвязанными величинами, и это осложняет задачу определения коэффи-

¹ Следует подчеркнуть, что согласование обеспечивает не только оптимальную передачу мощности, но и оптимальную передачу напряжения и тока. Поэтому вывод о целесообразности понижающего трансформатора в полной мере относится к усилителям *н а п р я ж е н и я*, в которых мощность как таковая не представляет интереса. Пусть, например, $R_H = 0,01 R_r$. Тогда в отсутствие трансформатора на нагрузке выделится примерно 0,01 напряжения источника сигнала; в случае согласования ($n = 0,1$) эта доля повышается до 0,05, а в случае повышающего трансформатора с $n = 10$ — уменьшается примерно до 0,001.

циента $n_{\text{согл}}$. В самом деле, коэффициент β_{oe} , который нужно использовать в формуле (7-2), есть функция сопротивления $R'_H = R_{\text{вых}}$, а коэффициент γ_6 , который нужно использовать в формуле (7-23), является функцией сопротивления $R'_r = R_{\text{вх}}$. Анализ, аналогичный тому, который приведен в § 11-3 применительно к коэффициенту усиления мощности (11-10), показывает, что в каскаде с последовательным включением трансформатора

$$n_{\text{согл}} = \sqrt{\frac{\beta(r_3 + r_6)}{r_k}}. \quad (11-46)$$

Поскольку $r_k \gg R_k$, коэффициент трансформации (11-46) значительно (примерно на порядок) меньше, чем в схеме с параллельным включением трансформатора (11-4а).

В заключение отметим, что трансформаторная связь часто используется не для повышения усиления, а только для преобразования сопротивлений согласно (11-1). Такое применение трансформаторов характерно для первого и последнего каскадов усилителя и объясняется тем, что ряд датчиков и некоторые виды нагрузки нуждаются для нормальной работы в определенном режиме, близком либо к холостому ходу, либо к короткому замыканию. При этом приходится мириться с тем, что коэффициент передачи K_u или K_i может оказаться меньше, чем в случае бестрансформаторной связи.

11-3. ОБЛАСТЬ СРЕДНИХ ЧАСТОТ

Трансформаторы могут включаться необязательно и на входе, и на выходе каскада. Но мы рассмотрим именно такой общий случай, поскольку в частных случаях можно в соответствующих формулах полагать $n_1 = 1$ или $n_2 = 1$. В основу анализа положим каскад ОЭ, хотя при трансформаторном соединении каскады ОБ также вполне работоспособны. Результаты анализа при соответствующей замене параметров (см. § 10-1) пригодны для обоих типов каскадов.

Параметры каскада. При анализе трансформаторных каскадов на средних частотах пренебрежем влиянием реактивных элементов, в том числе индуктивностями трансформаторов.

Для параллельной схемы, подставив в формулу (7-4а) пересчитанные сопротивления R'_r и R'_H и умножив правую часть на коэффициенты трансформации, получим:

$$K_u = \frac{\beta_{\text{oe}} (R_k \parallel R'_H)}{R'_r + R_{\text{вх}}} n_1 n_2. \quad (11-5а)$$

Аналогично из формулы (7-7а) получим:

$$K_i = \frac{\beta_{\text{oe}}}{n_1 n_2} \frac{R'_r}{R'_r + R_{\text{вх}}} \frac{R_k}{R_k + R'_H}, \quad (11-5б)$$

где $R'_r = R_r n_1^2$; $R'_H = R_H / n_2^2$, а коэффициент β_{oe} рассчитывается по формуле (7-20).

В случае последовательного включения трансформатора Tr_2 нужно в формулах (11-5) положить $R_k = \infty$; тогда

$$K_u = \frac{\beta_{oe} R'_H}{R'_r + R_{вх}} n_1 n_2; \quad (11-6a)$$

$$K_i = \frac{\beta_{oe}}{n_1 n_2} \frac{R'_r}{R'_r + R_{вх}}. \quad (11-6б)$$

Коэффициенты усиления в формулах (11-5) и (11-6) приняты положительными, так как их знаки зависят от включения обмоток.

Если источник сигнала и нагрузка не предъявляют особых требований соответственно к входному и выходному сопротивлениям, то, как указывалось выше, целесообразно согласовывать значения $R_{вх}$ и $R_{вых}$ с R_r и R_H , т. е. делать

$$n_1 = \sqrt{\frac{R_{вх}}{R_r}} \text{ и } n_2 = \sqrt{\frac{R_H}{R_{вых}}}.$$

Этот случай характерен для межкаскадной связи, когда источником сигнала служит выход предыдущего каскада, а нагрузкой — вход следующего каскада. При таком согласовании, когда $R'_r = R_{вх}$ и $R'_H = R_{вых}$, параметры (11-6) имеют максимальное значение:

$$K_{u \text{ согл}} = \frac{\beta_{oe}}{2} \sqrt{\frac{R_{вых}}{R_{вх}}} \sqrt{\frac{R_H}{R_r}}; \quad (11-7a)$$

$$K_{i \text{ согл}} = \frac{\beta_{oe}}{2} \sqrt{\frac{R_{вых}}{R_{вх}}} \sqrt{\frac{R_r}{R_H}}. \quad (11-7б)$$

Пусть, например, $R_H = R_{вх}$; $R_r = R_{вых} = 10 R_{вх}$ и $\beta_{oe} = 40$. Тогда для бестрансформаторного каскада, полагая $n_1 = n_2 = 1$ в формулах (11-6), получаем $K_u = 3,5$ и $K_i \approx 35$. Для согласованного каскада из формул (11-7) находим $K_u \approx 20$ и $K_i \approx 200$.

Максимальный коэффициент усиления мощности. Выведенные формулы позволяют найти предельное усиление мощности транзисторным каскадом в условиях трансформаторного согласования. Для этого нужно ориентироваться на последовательное включение выходного трансформатора, так как наличие сопротивления R_k в параллельном каскаде заведомо снижает усиление.

Напомним, что параметры K_u и K_i определялись как отношение напряжения или тока в нагрузке соответственно к э. д. с. E_r или току I_r источника сигнала. Коэффициент передачи мощности K_p принято определять как отношение мощности в нагрузке к мощности на входе четырехполюсника, т. е. после сопротивления R_r . В режиме согласования, когда $R_r = R_{вх}$, входное напряжение каскада составляет $1/2 E_r$, а входной ток $1/2 I_r$. Отсюда следует, что $P_{вх} = 1/4 E_r I_r$; значит,

$$K_{p \text{ согл}} = 4 K_{u \text{ согл}} K_{i \text{ согл}}. \quad (11-8)$$

Подставляя в (11-8) выражения (11-7), получаем:

$$K_{p \text{ согл}} = (\beta_{oe})^2 \frac{R_{\text{вых}}}{R_{\text{вх}}}. \quad (11-9)$$

Поскольку вход и выход каскада согласованы (т. е. $R'_Г = R_{\text{вх}}$ и $R'_Н = R_{\text{вых}}$), можно записать два уравнения:

$$R'_Г = (r_6 + r_9) (1 + \beta_{oe} \gamma_{60}); \quad R'_Н = r_K^* (1 + \beta \gamma_6),$$

где

$$\beta_{oe} = \beta \frac{r_K^*}{r_K^* + R'_Н}; \quad \gamma_6 = \frac{r_9}{r_9 + r_6 + R'_Г}.$$

Из этих двух уравнений получаются все необходимые параметры¹:

$$R'_Г = R_{\text{вх}} = (r_9 + r_6) \sqrt{1 + \beta \gamma_{60}};$$

$$R'_Н = R_{\text{вых}} = r_K^* \sqrt{1 + \beta \gamma_{60}};$$

$$\beta_{oe} = \frac{\beta}{1 + \sqrt{1 + \beta \gamma_{60}}}.$$

Теперь (11-9) можно записать в следующей форме:

$$K_{p \text{ согл}} = \frac{\beta^2 r_K^*}{(r_9 + r_6) (1 + \sqrt{1 + \beta \gamma_{60}})^2} = \frac{r_K}{r_9} \frac{\beta \gamma_{60}}{(1 + \sqrt{1 + \beta \gamma_{60}})^2}. \quad (11-10)$$

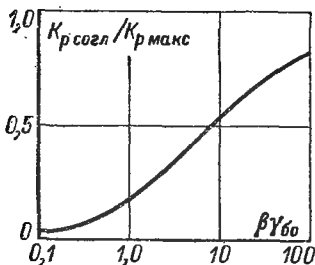
Формула (11-10) характеризует усилительную способность транзистора. Например, если $r_K = 1 \text{ МОм}$; $\beta = 40$; $r_9 = 25 \text{ Ом}$; $r_6 = 100 \text{ Ом}$, то $K_{p \text{ согл}} = 20 \text{ 000}$, или 43 дБ.

С уменьшением тока I_9 отношение r_K/r_9 меняется мало (см. § 4-5), а коэффициент γ_{60} растет. В пределе при $\gamma_{60} = 1$ и $\sqrt{\beta} \gg 1$ получаем:

$$K_{p \text{ макс}} \approx \frac{r_K}{r_9}.$$

Для приведенных выше параметров $K_{p \text{ макс}} \approx 40 \text{ 000}$, или 46 дБ.

Рис. 11-2. Зависимость коэффициента усиления мощности в режиме согласования от глубины внутренней обратной связи.



На рис. 11-2 показана зависимость $K_{p \text{ согл}}(\beta \gamma_{60})$, из которой видно, что практически достижимые коэффициенты усиления мощности при $\beta = 50 \div 200$ и $\gamma_{60} = 0,05 \div 0,5$ составляют 30—80% максимально возможных.

¹ Формула (11-46) получена путем подстановки приведенных ниже параметров $R_{\text{вх}}$ и $R_{\text{вых}}$ в выражение (11-3) с учетом соотношения (4-71).

В области средних частот мы пренебрегаем индуктивностями обмоток трансформаторов, считая их бесконечно большими. Между тем именно эти индуктивности наряду с емкостями обуславливают частотные искажения в области низших частот. Эквивалентные схемы входной и выходной цепей каскада для этой области показаны на рис. 11-3. Здесь собственные сопротивления обмоток включены в сопротивления R_r , $R'_{вх}$ и $R_{вых}$, R'_n , а все емкости (см. рис. 11-1) для простоты опущены.

Граничная частота и выбор индуктивности обмоток трансформатора. Наличие реактивного сопротивления X_{L_1} (рис. 11-3, а) приводит к уменьшению входного тока, а значит, и входного напряжения. Поэтому с уменьшением частоты параметры K_u и K_i должны уменьшаться. Их частотные характеристики получаются одинаковыми и имеют вид:

$$A = \frac{A_0}{1 + j \frac{\omega_n}{\omega}}, \quad (11-11)$$

где A_0 — значение параметра на средних частотах, а граничная частота выражается следующим образом¹:

$$\omega_{н1} = \frac{1}{\tau_{н1}} = \frac{R'_{вх} \parallel R_r}{L_1}. \quad (11-12)$$

Знак плюс в числителе (11-11) поставлен для общности. Вообще говоря, этот знак и соответственно фаза параметра зависят от включения обмоток. Нетрудно заметить, что граничная частота будет минимальной при $R_r \ll R'_{вх}$ и максимальной при $R_r \gg R'_{вх}$. При заданной граничной частоте (или постоянной времени) необходимое значение индуктивности первичной обмотки легко получается из (11-12). Например, если $f_n = 50$ Гц ($\omega_n \approx 300$ с⁻¹) и $R'_{вх} = R_r = 1$ кОм, то $L_1 \approx 2$ Г.

Перейдем к анализу выходной цепи. Ответвление тока эквивалентного генератора в индуктивность L_2 (рис. 11-3, б) приводит к уменьшению нагрузочного тока и выходного напряжения. Частотные характеристики параметров имеют вид (11-11), где граничная частота

$$\omega_{н2} = \frac{1}{\tau_{н2}} = \frac{R_{вых} \parallel R'_n}{L_2}. \quad (11-13)$$

В зависимости от значения R'_n граничная частота и постоянная времени будут меняться. Наилучшие частотные и переходные свой-

¹ В формуле (11-12) и на рис. 11-3, а не учтено эквивалентное сопротивление делителя в цепи базы; оно должно считаться подключенным параллельно входному сопротивлению.

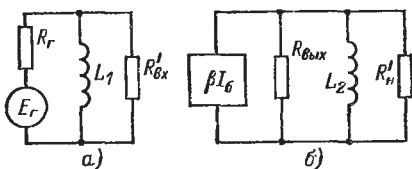


Рис. 11-3. Эквивалентные схемы входа (а) и выхода (б) трансформаторного каскада в области низших частот.

ства соответствуют малым значениям R'_H ; при больших R'_H частотные свойства ухудшаются.

Заметим, что при последовательном включении трансформатора Tr_2 выходное сопротивление (особенно в каскаде ОБ) велико и для больших R'_H (например, при наличии согласования) индуктивность L_2 нередко оказывается неприемлемо большой. Например, при $R_{\text{вых}} = R'_H = 50 \text{ кОм}$ и $f = 50 \text{ Гц}$ получается $L_2 \approx 100 \text{ Г}$. Тогда можно зашунтировать первичную обмотку сопротивлением R_K . Значение этого сопротивления находится из выражения

$$\tau_{H2} = \frac{L_2}{R_{\text{вых}} \parallel R'_H \parallel R_K},$$

в котором задаются приемлемым значением L_2 . Шунт R_K , конечно, на средних частотах снижает усиление каскада.

Искажения вершины импульса. Трансформаторные усилители сравнительно редко используются для усиления длинных импульсов. В тех случаях, когда это все же приходится делать, можно руководствоваться следующими соображениями.

Заменяв в выражении (11-11) оператор $j\omega$ оператором s , получим изображение переходной характеристики для одной из трансформаторных связей:

$$A(s) = A_0 \frac{s\tau_H}{1 + s\tau_H}. \quad (11-14)$$

Сама переходная характеристика будет обычной спадающей экспонентой. С учетом обеих трансформаторных связей изображение имеет вид $A_1(s) A_2(s)$, а с учетом влияния емкостей оно еще более усложняется. Однако, как и в случае усилителей с емкостной связью, знание переходной характеристики в течение большого интервала времени редко необходимо. Главный интерес представляет спад вершины при условии $t_H < t_{\text{сп}}$, где t_H — длительность импульса, $t_{\text{сп}}$ — время спада вершины [см. (7-32)]. При таком условии относительный спад вершины выражается формулой (7-31).

Следует заметить, что постоянные времени τ_{Hi} , входящие в (7-32а), вычисляются отдельно для каждой из реактивностей (C или L), имеющих на эквивалентной схеме, причем остальные реактивности полагаются при этом равными бесконечности. Например, постоянная времени емкости C_1 на рис. 11-1, а оказывается равной $C_1(R'_r + R_{\text{вх}} \parallel R_1 \parallel R_2)$, а постоянная времени емкости C_2 определяется формулой (7-30а), если положить $R_r = R'_r \parallel R_1 \parallel R_2$.

При оценке результирующей граничной частоты каскада можно пользоваться формулой (7-33а), где частичные граничные частоты соответствуют как индуктивностям обмоток [см. (11-12), (11-13)], так и блокирующим емкостям.

11-5. МАКСИМАЛЬНАЯ ЧАСТОТА ГЕНЕРАЦИИ ТРАНЗИСТОРА

Максимальная частота генерации не является параметром, специфичным только для генераторов¹. Она характеризует предельные частотные и усилительные возможности транзисторов независимо от схемы использования их. Поэтому этот важнейший параметр рассматривается в настоящем параграфе, исходя из теории трансформаторных усилителей.

¹ Генераторы синусоидального напряжения в данной книге не рассматриваются; им посвящена обширная литература (см., например, [67, 69, 133]).

Очевидно, что самовозбуждение схемы (при наличии обратной связи) возможно только тогда, когда коэффициент передачи мощности в каскаде превышает единицу. Поскольку коэффициент K_p является, вообще говоря, функцией частоты, максимальную частоту генерации (или, что то же самое, *максимальную частоту усиления мощности*) можно найти из условия $K_p(\omega) = 1$. При этом подразумевается, что в каскаде выполнены условия согласования, так как в противном случае передача мощности не будет оптимальной и соответственно максимальная частота не будет предельно возможной¹.

Положим в основу анализа выражение (11-10), но учтем при этом частотную зависимость параметров. Вместо r_k используем величину

$$|Z_k| = |r_k| \left| \left(\frac{1}{j\omega C_k} \right) \right| \approx \frac{1}{\omega C_k}$$

(поскольку рассматривается область весьма высоких частот), а для коэффициента передачи β примем:

$$\beta = \left| \frac{\beta}{1 + j\omega/\omega_\beta} \right| \approx \left| \frac{\beta\omega_\beta}{j\omega} \right| \approx \frac{\omega_\beta}{\omega}.$$

Кроме того, положим $|\beta| \gamma_{60} \ll 1$. Это вполне оправдано не только тем, что β существенно уменьшается с ростом частоты, но и тем, что сопротивление r_3 (а следовательно, и коэффициент γ_{60}), как увидим ниже, выгодно делать как можно меньше, чтобы повысить максимальную частоту генерации.

Подставляя в выражение (11-10) значения $|Z_k|$, $|\beta|$ и $\sqrt{1 + |\beta| \gamma_{60}} \approx 1$, получаем коэффициент передачи мощности в области высоких частот:

$$K_p = \frac{\omega_\beta}{4\omega^2 C_k (r_3 + r_6)}. \quad (11-15)$$

Теперь, полагая $K_p = 1$, легко найти максимальную частоту генерации:

$$f_{\text{ген}} \approx \sqrt{\frac{f_\beta}{8\pi (r_3 + r_6) C_k}}.$$

Как видим, величина $f_{\text{ген}}$ растет с уменьшением сопротивления r_3 , о чем говорилось выше. При условии $r_3 \ll r_6$, т. е. при достаточно большом эмиттерном токе, частота $f_{\text{ген}}$ достигает предельного значения:

$$f_{\text{ген. макс}} \approx \sqrt{\frac{f_\beta}{8\pi r_6 C_k}}. \quad (11-16)$$

¹ На высоких частотах, когда реактивные составляющие импедансов сравнимы с активными, согласование подразумевает не только выполнение условия (11-2), но и компенсацию реактивных составляющих, т. е. обеспечение резонансов во входной и выходной цепях.

Из выражения (11-16) ясно, что максимальная частота генерации является более универсальным параметром, чем отдельные постоянные времени τ_α и $\tau_\beta = r_\beta C_k$. Поэтому величина $f_{\text{ген. макс}}$ часто приводится в справочниках, тем более что в случае высокочастотных транзисторов она измеряется проще, чем f_α . Частота $f_{\text{ген. макс}}$ может быть больше или меньше, чем f_α , в зависимости от соотношения величин ω_α и $1/\tau_\beta$. Обычно у низкочастотных бездрейфовых транзисторов $f_{\text{ген. макс}} > f_\alpha$, а у высокочастотных дрейфовых транзисторов $f_{\text{ген. макс}} < f_\alpha$.

Глава двенадцатая

МОЩНЫЕ ВЫХОДНЫЕ КАСКАДЫ

12-1. ВВЕДЕНИЕ

Усиление мощности есть необходимое свойство всякого усилителя. Поэтому название «усилитель мощности» в применении к выходным каскадам не отражает их специфики. Под мощным каскадом мы будем понимать такой усилитель, для которого заданы величина нагрузки R_n и мощность в этой нагрузке P_n .

Коэффициент усиления мощности K_p обычно не задается, хотя его желательно иметь большим. Рассчитав выходной каскад по заданным значениям P_n и R_n , затем находят значение K_p , чтобы оценить мощность, которую должен давать предыдущий каскад.

Поскольку значительная выходная мощность связана с затратой большой мощности источником питания, важное значение приобретает такой параметр выходного каскада, как к. п. д.

Большая мощность означает работу с большими токами и напряжениями. Поэтому в мощных каскадах переменные составляющие токов и напряжений сравнимы с постоянными составляющими. При этом проявляется зависимость основных параметров транзисторов α или β от режима, а также нелинейность входных характеристик $I_g(U_g)$ или $I_b(U_b)$. Отсюда вытекает необходимость оценки нелинейных искажений усилителя. В случае синусоидального сигнала (который всегда принимается при анализе мощных каскадов) такая оценка дается с помощью коэффициента нелинейных искажений — клирфактора [62, стр. 37]:

$$K_{\text{н. и}} = \frac{\sqrt{\sum_{i=2}^{\infty} I_{mi}^2}}{I_{m1}},$$

где I_{m1} — амплитуда 1-й (основной) гармоники, а I_{mi} — амплитуды высших гармоник, обуславливающих нелинейные искажения.

Амплитуды гармоник можно найти графоаналитическими методами с использованием «проходной» характеристики $I_{\text{вых}}(U_{\text{вх}})$ (см. [134], § 4-2). Так, если выходной ток (при синусоидальном вход-

ном. напряжении) содержит в основном 2-ю гармонику, т. е. кривая $i_{\text{вых}}(t)$ асимметрична, то так называемый *метод трех ординат* дает коэффициент нелинейных искажений в виде

$$K_{\text{н. и}} = \frac{I_{\text{макс}} + I_{\text{мин}} - 2I_0}{2(I_{\text{макс}} - I_{\text{мин}})}, \quad (12-1)$$

где $I_{\text{макс}}$ и $I_{\text{мин}}$ — максимальное и минимальное значения выходного тока за период, а I_0 — значение тока в момент, когда $U_{\text{вх}} = 0$. В том случае, когда существенную роль играет 3-я гармоника, т. е. когда кривая $i_{\text{вых}}(t)$ симметрична относительно оси времени, применяется несколько более громоздкий *метод пяти ординат* [62, § 2-3].

Транзисторные выходные каскады, как и ламповые, могут выполняться *однотактными* и *двухтактными*, а транзисторы в них могут работать в *классах А, В и АВ* (см. [62], § 2-4, 4-3). Ниже рассматриваются наиболее типичные схемы выходных каскадов — однотактный каскад класса А и двухтактный каскад класса В. Поскольку в мощных усилителях часто используется трансформаторная связь, то наряду с каскадами ОЭ широкое распространение имеют каскады ОБ. В принципе можно использовать и каскады ОК, но, будучи повторителями, они требуют для раскачки больших входных напряжений и в этом отношении уступают каскадам ОЭ: получается значительно меньший коэффициент усиления мощности. Поэтому включение ОК используется лишь в специальных схемах, например в каскадах с дополнительной симметрией (см. § 12-3).

12-2. ОДНОТАКТНЫЕ КАСКАДЫ КЛАССА А

Последующий анализ проводится подробно для каскада ОБ как более простого, а относительно каскада ОЭ делаются лишь необходимые замечания, поскольку основные соотношения и выводы оказываются общими.

Энергетические соотношения. Схема каскада, подлежащая анализу, показана на рис. 12-1, а коллекторное семейство характеристик транзистора — на рис. 12-2. Точка покоя А (в отсутствие сигнала) расположена на линии *статической* нагрузки r_1 , проведенной из абсциссы $E_{\text{к}}$. Сопротивление r_1 есть сопротивление первичной обмотки выходного трансформатора. Обычно оно очень мало, так что линия r_1 идет почти вертикально и можно считать $U_{\text{кА}} = E_{\text{к}}$. Эмиттерный ток покоя $I_{\text{эА}}$ задается цепью $E_{\text{э}}, R_{\text{э}}$.

Входной сигнал меняет ток и перемещает рабочую точку транзистора по линии *динамической* нагрузки $R'_{\text{н}}$, проходящей через точку А. Сопротивление $R'_{\text{н}}$ есть пересчитанное к первичной обмотке сопротивление истинной нагрузки $R_{\text{н}}$. Мощность в пересчитанной нагрузке выражается формулами

$$P'_{\text{н}} = \frac{1}{2} U_{\text{км}} I_{\text{км}} = \frac{1}{2} I_{\text{км}}^2 R'_{\text{н}} = \frac{1}{2} U_{\text{км}}^2 / R'_{\text{н}}. \quad (12-2)$$

Мощность в истинной нагрузке связана с величиной P'_H соотношением

$$P_H = \eta_{\text{тр}} P'_H,$$

где $\eta_{\text{тр}}$ — к. п. д. выходного трансформатора. Зная мощность P_H и задаваясь значением $\eta_{\text{тр}} = 0,75 \div 0,95$ (тем большим, чем больше мощность), находят необходимое значение P'_H .

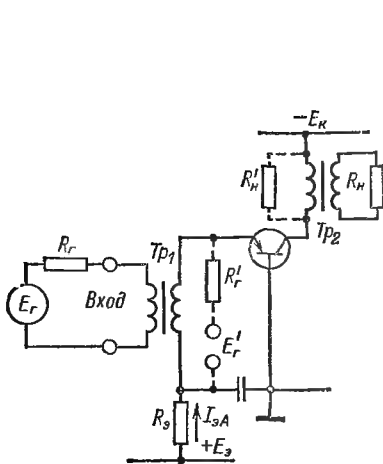


Рис. 12-1. Однотактный каскад ОБ класса А.

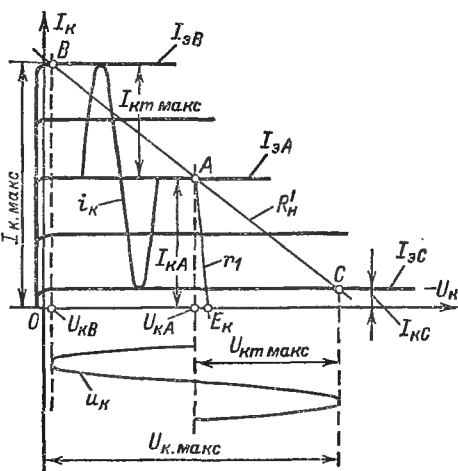


Рис. 12-2. Режим работы транзистора в каскаде ОБ класса А.

Для наиболее полного использования транзистора координаты точки покоя должны удовлетворять соотношениям

$$U_{KA} \approx U_{K \text{ макс}}; \quad (12-3a)$$

$$I_{KA} \approx I_{K \text{ макс}}; \quad (12-3б)$$

где $I_{K \text{ макс}}$ и $U_{K \text{ макс}}$ — максимальные амплитуды напряжения и тока, соответствующие максимальному входному сигналу. При таком выборе точки А (рис. 12-2) максимальному току $I_{K \text{ макс}}$ соответствует минимальное коллекторное напряжение U_{KB} , а максимальному напряжению $U_{K \text{ макс}}$ — минимальный коллекторный ток I_{KC} . Из рис. 12-2 получаем зависимости:

$$U_{K \text{ макс}} = 2U_{K \text{ макс}} + U_{KB}; \quad (12-4a)$$

$$I_{K \text{ макс}} = 2I_{K \text{ макс}} + I_{KC}. \quad (12-4б)$$

Значения $U_{K \text{ макс}}$ и $I_{K \text{ макс}}$ не должны превышать соответствующих допустимых значений. Тогда, полагая $U_{KB} \ll U_{K \text{ доп}}$ и

$I_{кс} \ll I_{к. доп}$, что всегда оправдано, приходим к следующим ограничениям для максимальных амплитуд¹:

$$U_{к т макс} < \frac{1}{2} U_{к. доп}; \quad (12-5a)$$

$$I_{к т макс} < \frac{1}{2} I_{к. доп}. \quad (12-56)$$

Руководствуясь соотношениями (12-5), а также формулами (12-1), выбирают максимальные амплитуды для получения необходимой мощности $P'_{н. макс}$. После этого сопротивление $R'_н$ можно найти из формулы

$$R'_н = \frac{U_{к т макс}}{I_{к т макс}}, \quad (12-6)$$

а необходимый коэффициент трансформации выходного трансформатора — из формулы

$$n_2 = \sqrt{R_н / R'_н}. \quad (12-7)$$

Мощность, отдаваемая источником питания, равна:

$$P_0 = E_к I_{кА} \quad (12-8)$$

и не зависит от значения сигнала. Разделив (12-2) на (12-8), получим к. п. д. коллекторной цепи:

$$\eta_к = \frac{1}{2} \frac{U_{к т} I_{к т}}{E_к I_{кА}}. \quad (12-9)$$

При максимальном сигнале и соблюдении условий $I_{к т макс} \approx I_{кА}$ и $U_{к т макс} \approx E_к$ (т. е. $I_{кс} \ll I_{к т макс}$ и $I_{кА} r_1 + U_{кВ} \ll U_{к т макс}$) максимальный к. п. д. $\eta_к$ оказывается близким к предельному значению 50% и практически может составлять до 40—45%. Общий к. п. д. каскада η_0 [с учетом потерь в выходном трансформаторе и особенно в цепи смещения ($E_э$, $R_э$)] значительно меньше указанной величины: обычно $\eta_0 = 30 \div 35\%$.

Перейдем к оценке максимальной мощности $P'_{н. макс}$. Перемножив правые и левые части выражений (12-5) и учитывая (12-2), легко получить:

$$P'_{н. макс} < \frac{1}{8} U_{к. доп} I_{к. доп}. \quad (12-10)$$

Однако максимальная мощность ограничена не только допустимыми током и напряжением, но и допустимой мощностью рассеяния на коллекторе $P_{к. доп}$, причем это ограничение в ряде слу-

¹ Следует напомнить, что допустимый ток транзистора ограничен не только и не только от допустимого напряжения и допустимой мощности и что совмещение предельных режимов (в данном случае по напряжению и току) не разрешается из соображений надежности. Иначе говоря, оба неравенства (12-5) должны выполняться с достаточным запасом (на 20—30%). Значения $U_{к. доп}$ и $I_{к. доп}$ должны соответствовать максимальной рабочей температуре.

чаев может быть главным. Из соотношения

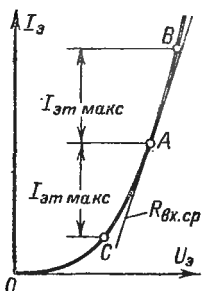
$$P_k = P_0 - P'_n$$

следует, что максимальная мощность P_k рассеивается в режиме покоя (при $P'_n = 0$) и равна P_0 . Выражая мощность P_0 как P'_n/η_k , получаем ограничение:

$$P'_{n, \text{ макс}} < \eta_k P_{k, \text{ доп}}. \quad (12-11)$$

Подбирая транзистор по заданной мощности в нагрузке или оценивая возможности заданного транзистора, нужно использовать обе формулы (12-10) и (12-11), учитывая при этом замечания в сноске на с. 415.

Полный расчет каскада включает в себя определение входной мощности $P_{вх}$, необходимой для раскачки каскада. В общем случае расчет величины $P_{вх}$ осложняется тем, что входная характеристика $I_g(U_g)$ нелинейна. Однако, если рабочий диапазон эмиттерных токов лежит в пределах ее омического участка, последний можно линеаризовать и использовать в расчетах усредненную величину входного сопротивления $R_{вх.ср}$ (рис. 12-3, где точки A, B и C соответствуют этим же точкам на рис. 12-2). Тогда



$$P_{вх} \approx \frac{I_{э м}^2 R_{вх.ср}}{2}. \quad (12-12)$$

Рис. 12-3. Входная характеристика каскада ОБ и ее идеализация для расчета входной мощности.

На омическом участке ток достаточно велик, поэтому можно пренебречь сопротивлением эмиттерного перехода и получить из формулы (10-1):

$$R_{вх.ср} \approx (1 - \alpha) r_6 = \frac{r_6}{1 + \beta}.$$

Разделив (12-1) на (12-12), найдем коэффициент усиления мощности; полагая $I_{эм} = I_{км}$, запишем его в следующих трех формах:

$$K_p = \frac{R'_n}{R_{вх.ср}} \approx \frac{R'_n}{r_6} (1 + \beta) \approx S R'_n. \quad (12-13)$$

Первая форма удобна при графической линеаризации входной характеристики, вторая — в случае транзисторов малой и средней мощности (для которых указывается коэффициент β , а сопротивление r_6 лежит в пределах от 10—20 до 1—2 Ом), третья — в случае транзисторов средней и большой мощности (для которых обычно указывается крутизна S). В каскаде ОБ коэффициент усиления получается невысоким (20—30, а иногда и меньше).

Выражения (12-13) показывают, что сопротивление R'_n желательно делать большим. Поэтому, как следует из (12-2), заданную мощность целесообразно обеспечивать повышением амплитуды n а п р я ж е н и я, а не тока.

Нелинейные искажения. Транзисторный усилитель характеризуется тем, что нелинейные искажения возникают не только в выходной (коллекторной), но и во входной цепи. В первом случае искажения обусловлены зависимостью коэффициента α от тока, во втором — нелинейностью входной характеристики.

Пусть задан синусоидальный ток i_s , источником которого служит генератор тока (рис. 12-4, а). Тогда кривая $u_s(t)$ будет сильно искажена, но это не имеет значения, так как параметром выходного семейства (см. рис. 12-2) является ток I_s . Кривая $i_k(t) = \alpha i_s$ будет почти синусоидальной, хотя в принципе положительная полуволна должна быть несколько меньше отрицательной из-за уменьшения коэффициента α в области больших токов.

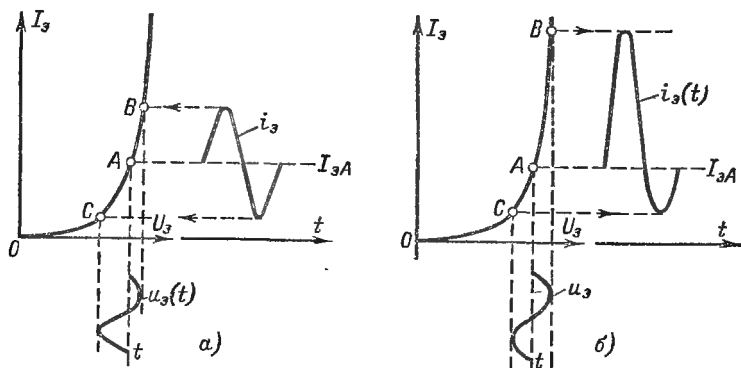


Рис. 12-4. Нелинейные искажения во входной цепи каскада ОБ.

а — при заданном токе; б — при заданном напряжении.

Пусть теперь задано синусоидальное напряжение u_s , источником которого служит генератор э. д. с. (рис. 12-4, б). Тогда входной ток $i_s(t)$, а вместе с ним и ток $i_k(t)$ окажутся резко несинусоидальными: будет сильно «приплюснута» отрицательная полуволна. С увеличением пересчитанного сопротивления источника сигнала R'_r нелинейные искажения уменьшаются, так как при этом режим входной цепи приближается к режиму заданного эмиттерного тока. Однако с увеличением сопротивления R'_r возрастают потери мощности в нем, т. е. ухудшается передача мощности от предоконечного каскада. Поэтому значения $R'_r \gg R_{\text{вх.ср}}$ невыгодны. Из рис. 12-4 ясно, что нелинейные искажения будут меньше, если не используется начальный криволинейный участок входной характеристики. Значит, минимальный эмиттерный ток $I_{\text{эс}}$ следует выбирать достаточно большим (у мощных транзисторов — десятки и сотни миллиампер), хотя при этом к. п. д. коллекторной цепи будет несколько снижаться.

Методика расчета нелинейных искажений и выбора сопротивления R'_r иллюстрируется на рис. 12-5. Сначала выбирают точки С и В. Первой соответствуют токи $I_{\text{эс}}$ (см. выше) и $I_{\text{к.мин}}$, второй —

ток $I_{к. макс} = 2I_{кт} + I_{к. мин}$. Затем задаются сопротивлением $R'_r \lesssim R_{вх. ср}$ и проводят из точек B и C линии нагрузки R'_r до пересечения с осью U_3 ; на ней получают точки b и c , определяющие двой-

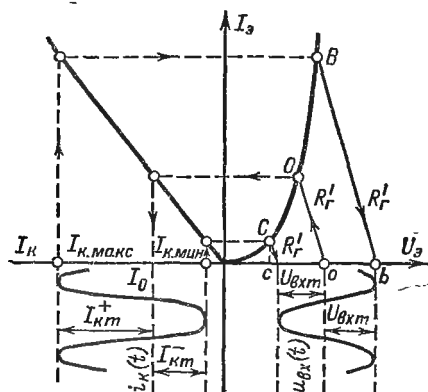


Рис. 12-5. Графоаналитический расчет нелинейных искажений (метод трех ординат). Исходные величины $I_{к. макс}$ и $I_{к. мин}$; искомая величина I_0 (определяется по точке o , лежащей посередине между b и c).

ную амплитуду напряжения $U_{вх}$. Точка o , лежащая посередине между b и c , определяет ток I_0 . Если коэффициент $K_{н.н}$, вычисленный по формуле (12-1), не удовлетворяет заданным требованиям, следует увеличить R'_r или $I_{э. мин}$ и повторить расчет.

Коэффициент трансформации входного трансформатора определяется из очевидного соотношения

$$n_1 = \sqrt{R'_r / R_r}, \quad (12-14)$$

в котором величина R'_r считается известной либо из расчета нелинейных искажений, либо из условия согласования $R'_r \approx R_{вх}$.

Свойства мощных каскадов в области низких частот такие же, как у маломощных каскадов (см. § 11-4). Поэтому выбор индуктивностей трансформаторных обмоток можно проводить, исходя из формул (11-12) и (11-13).

Особенности каскада ОЭ. Схема мощного каскада, показанная на рис. 12-6, отличается от схемы маломощного трансформаторного каскада (см. рис. 11-1, а) прежде всего отсутствием сопротивления R_3 . Такое отличие объясняется тем, что при токе эмиттера от 0,1 А и выше сопротивление R_3 , а значит, и сопротивление делителя в цепи базы должны были бы иметь очень малое значение. Это привело бы к резкому увеличению расхода мощности. Кроме того, конденсатор C_3 должен был бы иметь огромную емкость. Использование диода в нижнем плече базового делителя удобно тем, что при этом легко получить необходимое смещение на базе и в то же время обеспечить очень малое дифференциальное сопротивление нижнего плеча, т. е. достаточную стабилизацию режима. Кроме того, температурной стабильности способствуют близкие по величине температурные коэффициенты напряжений на диоде и на эмиттерном переходе транзистора.

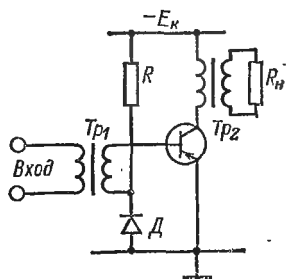


Рис. 12-6. Однотактный каскад ОЭ класса А.

Главное преимущество каскада ОЭ перед каскадом ОБ заключается в меньшей входной мощности; иначе говоря, каскад ОЭ имеет больший коэффициент усиления мощности. Этот выигрыш, казалось бы, можно оценить величиной $1 + \beta$, поскольку базовый ток в $1 + \beta$ раз меньше эмиттерного. Однако в схеме ОЭ допустимое коллекторное напряжение ниже, чем в схеме ОБ (см. § 4-7), а это ограничивает величину R'_H . Поэтому согласно (12-13) выигрыш в коэффициенте усиления мощности снижается по сравнению с величиной $1 + \beta$, которая подразумевает одинаковые значения R'_H в обеих схемах. Тем не менее выигрыш в значении K_p на порядок можно считать типичным.

По другим показателям каскад ОЭ уступает каскаду ОБ. В этом легко убедиться из рис. 12-7, где точки А, В и С по-прежнему соответствуют режиму покоя, минимальному напряжению и минимальному току.

Прежде всего в схеме ОЭ рабочий диапазон напряжений используется хуже, чем в схеме ОБ, так как напряжение в точке В принципиально не равно нулю. Соответственно меньше оказывается максимальный к. п. д. коллекторной цепи. Температурная зависимость $\beta(T)$ и отсутствие сопротивления R_e несколько ухудшают стабильность рабочей точки и повышают температурную зависимость $\eta_k(T)$. Наконец, каскад ОЭ дает большие нелинейные искажения в коллекторной цепи: они обусловлены ярко выраженной зависимостью β от тока. В связи с этим амплитуды токов и напряжений во время положительного и отрицательного полупериодов заметно различаются (на рис. 12-7 эти амплитуды обозначены знаками «+» и «—»)¹.

Расчет нелинейных искажений осуществляется тем же графоаналитическим методом, который описан выше. Следует заметить, что искажения, возникающие в выходной цепи, могут частично компенсировать искажения, возникающие во входной цепи, так как в последнем случае подчеркивается положительная полуволна

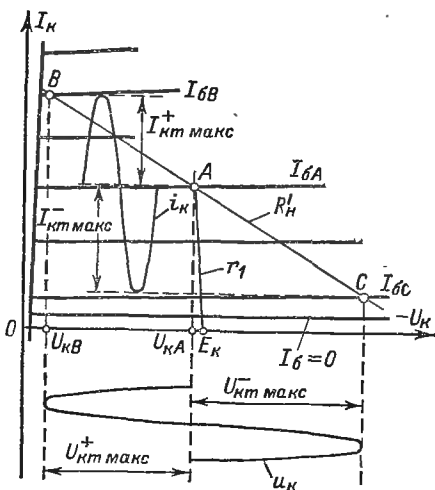


Рис. 12-7. Режим работы транзистора в каскаде ОЭ класса А.

¹ Строго говоря, амплитуды $I_{кт макс}^+$ и $I_{кт макс}^-$ следует отсчитывать от уровня I_0 , который не совпадает с уровнем $I_{кА}$. Однако, поскольку указанные уровни различаются не очень сильно, на рис. 12-5 и 12-7 они приняты совпадающими.

тока (рис. 12-4, б), а в первом — отрицательная (рис. 12-7). Следовательно, при некотором оптимальном сопротивлении R'_r можно добиться условия

$$\beta^+ I_{0\text{ т макс}}^+ = \beta^- I_{0\text{ т макс}}^-$$

при котором $I_{\text{к т макс}}^+ = I_{\text{к т макс}}^-$. В этом случае 2-я и вообще четные гармоники оказываются несущественными и рассчитывать коэффициент нелинейных искажений по формуле (12-1) нельзя; нужно использовать метод пяти ординат [62, 134]¹. Однако на практике оптимальное значение $R'_{\text{г. опт}}$ нередко оказывается значительно большим значения $R_{\text{вх. ср}}$, а это означает, что при реализации оптимального сопротивления $R'_{\text{г. опт}}$ минимум нелинейных искажений будет покупаться ценой рассогласования данного каскада с предыдущим, т. е. ценой ухудшения общего коэффициента усиления мощности.

Коэффициент трансформации n_1 определяется формулой (12-15), в которой под R'_r обычно понимают $R_{\text{вх. ср}}$ или несколько большую величину (в последнем случае достигается компромисс между условием согласования и минимальными нелинейными искажениями).

12-3. ДВУХТАКТНЫЕ КАСКАДЫ КЛАССА В

Для транзисторных усилителей мощности более характерна работа в классе В, чем в классе А, в связи со значительно большей

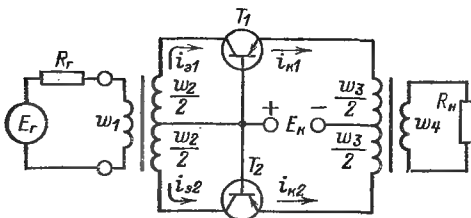


Рис. 12-8. Двухтактный каскад ОБ класса В.

экономичностью первого. В основу анализа кладется каскад ОБ, потому что каскад ОЭ имеет те же относительные преимущества и недостатки, которые рассмотрены в § 12-2.

Энергетические соотношения. Схема каскада, приведенная на рис. 12-8, отличается от аналогичной ламповой схемы отсутствием источника смещения. Такой источник не нужен, так как каждый из транзисторов запирается в самом начале соответствующего отрицательного полупериода (для кремниевых транзисторов — в конце положительного полупериода, см. входную характеристику на рис. 2-21, б). При поступлении синусоидального сигнала оба транзистора работают поочередно, каждый в течение одного полупериода (рис. 12-9 и 12-10); в нагрузке токи от обоих транзисторов складываются.

¹ Анализ показывает, что даже при оптимальном значении R'_r нелинейные искажения в каскаде ОЭ больше, чем в каскаде ОБ, при одной и той же выходной мощности.

В отсутствие сигнала рабочая точка транзистора (рис. 12-9) определяется линией нагрузки r_1 и кривой, соответствующей очень малому току I_0 , протекающему при коротком замыкании эмиттера с базой. Практически можно считать $I_{KA} = I_{K0}$ и $U_{KA} = E_K$. Для полного использования транзистора нужно, как и в одноконтном

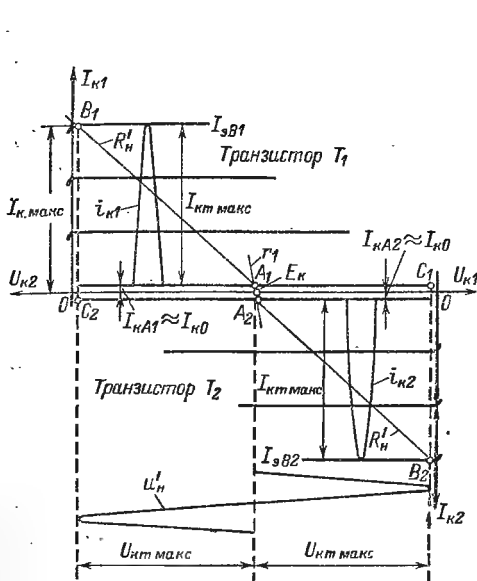


Рис. 12-9. Режим работы транзистора в каскаде ОБ класса В.

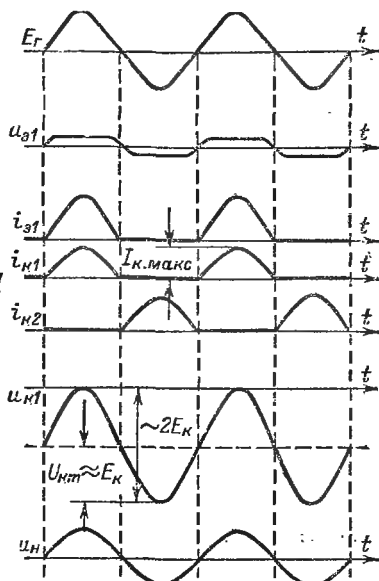


Рис. 12-10. Временные диаграммы токов и напряжений в двухтактном каскаде класса В.

каскаде, выполнить соотношение $U_{KA} \approx U_{K \text{ max}}$. Вместо выражений (12-4) получаются выражения

$$U_{K \text{ max}} = 2U_{K \text{ max}} + U_{KB}; \quad (12-15a)$$

$$I_{K \text{ max}} = I_{K \text{ max}} + I_{KA}; \quad (12-15b)$$

а вместо неравенств (12-5) — неравенства

$$U_{K \text{ max}} < \frac{1}{2} U_{K \text{ доп}}; \quad (12-16a)$$

$$I_{K \text{ max}} < I_{K \text{ доп}}. \quad (12-16b)$$

Руководствуясь соотношениями (12-16), а также формулами (12-2), выбирают максимальные амплитуды для получения необходимой мощности $P'_{n \text{ max}}$. После этого сопротивление R'_n определяется из формулы (12-6), а коэффициент трансформации $n_2 = \omega_4 / \left(\frac{1}{2} \omega_3 \right)$ — из формулы (12-7)¹.

¹ При использовании неравенств (12-16) следует учитывать необходимый «запас», оговоренный в сноске на с. 415.

Средний ток через каждый транзистор равен $\frac{1}{\pi} I_{км} + I_{кА}$, а следовательно, средний ток через источник питания

$$I_{0\text{ ср}} = 2 \left(\frac{1}{\pi} I_{км} + I_{кА} \right); \quad (12-17)$$

поэтому мощность, отбираемая от источника питания, выразится следующим образом:

$$P_0 = E_{к} I_{0\text{ ср}} = \frac{2}{\pi} E_{к} (I_{км} + \pi I_{кА}). \quad (12-18)$$

Деля (12-2) на (12-18), получаем к. п. д. коллекторной цепи:

$$\eta_{к} = \frac{\pi}{4} \frac{U_{км}}{E_{к}} \frac{I_{км}}{I_{км} + \pi I_{кА}}. \quad (12-19)$$

При максимальном сигнале и соблюдении условий $U_{км\text{ макс}} \approx E_{к}$ и $I_{км\text{ макс}} \gg \pi I_{кА}$ максимальный к. п. д. близок к предельному значению 78% и практически может составить 65—70%, конечно, без учета потерь в трансформаторе.

Перемножая правые и левые части выражений (12-16) и учитывая (12-2), легко получить:

$$P'_{н. \text{ макс}} < \frac{1}{4} U_{к. \text{ доп}} I_{к. \text{ доп}}. \quad (12-20)$$

Эта мощность, естественно, вдвое больше, чем (12-10), поскольку в двухтактном каскаде используются два транзистора. Теперь оценим максимальную мощность с точки зрения допустимой мощности рассеяния на коллекторе. Для каждого транзистора имеем:

$$P_{к} = \frac{1}{2} (P_0 - P'_{н}).$$

В классе В от величины сигнала (т. е. от тока $I_{км}$) зависит не только мощность $P'_{н}$, но и мощность P_0 . Подставив в выражение для $P_{к}$ формулы (12-18) и (12-2) (последнюю в виде $P'_{н} = \frac{1}{2} I_{км}^2 R'_{н}$ и продифференцировав $P_{к}$ по току $I_{км}$, нетрудно получить амплитуду тока, обеспечивающую максимальную мощность рассеяния: $I_{км} = \frac{2}{\pi} E_{к} / R'_{н}$.

При такой амплитуде имеем:

$$P_{к. \text{ макс}} = \frac{1}{\pi^2} E_{к}^2 / R'_{н}. \quad (12-21)$$

Если принять $E_{к} \approx U_{км\text{ макс}}$ и положить $P_{к. \text{ макс}} < P_{к. \text{ доп}}$, то, учитывая (12-2), приходим к соотношению

$$P'_{н. \text{ макс}} < \frac{\pi^2}{2} P_{к. \text{ доп}} \approx 5 P_{к. \text{ доп}}. \quad (12-22)$$

Это средняя мощность за период. В течение периода мгновенная мощность меняется, и в связи с небольшой тепловой инерцией транзисторов $p-n$ переходы могут перегреваться во время пиков мгновенной мощности несмотря на допустимую среднюю температуру [135]. Для того чтобы учесть это специфическое обстоятельство, примем опять $U_{к т макс} = E_k$ (рис. 12-10) и запишем выражение для мгновенной мощности рассеяния на одном из коллекторов при максимальном токе:

$$p_k = u_k i_k \approx E_k (1 - \cos \omega t) I_{к т макс} \cos \omega t.$$

Эта функция имеет два максимума в течение положительного полупериода, когда протекает ток i_k . Максимумы получаются при $\omega t = \pm \pi/3$ и составляют:

$$p_{к. макс} = 0,25 E_k I_{к т макс} \approx 0,5 P'_{н. макс}. \quad (12-23)$$

Если ограничить значение $p_{к. макс}$ допустимой мощностью $P_{к. доп}$, то это будет служить полной гарантией от перегрева. Однако такая гарантия излишня, так как транзистор все же обладает тепловой инерцией и температура перехода всегда ниже той, которая соответствует пику мощности $p_{к. макс}$, даже при низкой частоте. Поэтому имеет смысл уменьшить коэффициент в правой части (12-22), но не до 2, как следует из (12-23), а до компромиссного значения 3 и для расчетов принять:

$$P'_{н. макс} < 3 P_{к. доп}. \quad (12-24)$$

Дальнейшее увеличение выходной мощности достигается параллельным включением транзисторов в плечах каскада. При этом необходимо обеспечить равномерное распределение токов (а значит, и мощностей) между транзисторами. Для этого в цепи эмиттеров включают добавочные сопротивления $R_э$, которые повышают величину $R_{вх}$ и тем самым уменьшают входной и коллекторный токи. Очевидно, что добавочное сопротивление должно быть тем больше, чем больше ток данного транзистора или (при одинаковых входных напряжениях) чем больше его крутизна S .

При выборе значения $R_э$ руководствуются одной из двух предпосылок [136] — либо обеспечением одинаковых эквивалентных крутизн всех транзисторов (путем индивидуального подбора $R_э$), либо включением сопротивлений $R_э$, достаточно больших, чтобы в нужной степени выровнять эквивалентные крутизны.

Первый метод приводит к следующим результатам. Если в выражение для крутизны транзистора (4-95а) подставить сумму $R_э + r_э$ вместо $r_э$, поделить числитель и знаменатель на α и положить коэффициент $1/\alpha$ при $R_э$ равным единице, то эквивалентная крутизна может быть записана в виде

$$S_{эkv} = \frac{1}{R_э + 1/S},$$

где S — крутизна транзистора (без $R_э$). В цепь транзистора с минимальной крутизной $S_{мин}$ сопротивление можно не включать, а в цепи остальных транзисторов следует включить такие сопротивления, при которых выполняется условие $S_{эkv} = S_{мин}$. Для i -го транзистора

$$R_{эi} = \frac{1}{S_{мин}} - \frac{1}{S_i}. \quad (12-25)$$

При использовании третьего метода исходят из того, что отношение коллекторных токов равно отношению крутизн. Следовательно, если $R_3 = 0$, то для транзисторов с максимальным и минимальным токами имеем:

$$\delta I = I_{\text{макс}}/I_{\text{мин}} = S_{\text{макс}}/S_{\text{мин}}.$$

При наличии одинаковых сопротивлений R_3 отношение токов определяется эквивалентными крутизнами, т. е.

$$\delta I' = I'_{\text{макс}}/I'_{\text{мин}} = \frac{R_3 + 1/S_{\text{мин}}}{R_3 + 1/S_{\text{макс}}} = \frac{R_3 S_{\text{макс}} + \delta I}{R_3 S_{\text{макс}} + 1}.$$

Если значение δI известно, а значение $\delta I'$ задано, то из последнего выражения легко найти необходимую величину добавочных сопротивлений:

$$R_3 = \frac{1}{S_{\text{макс}}} \frac{\delta I - \delta I'}{\delta I' - 1}. \quad (12-26)$$

Второй метод, как правило, предпочтителен, так как он не связан с индивидуальным подбором сопротивлений. Такой подбор тем более затруднителен, что значение R_3 обычно не превышает десятых долей ома.

Вопрос о входной мощности в двухтактном каскаде решается на основе суммарных (совмещенных) входных характеристик (рис. 12-11). Идеализация этих характеристик, показанная пунктиром, позволяет оценить входную мощность по весьма приближенной формуле:

$$P_{\text{вх}} \approx \frac{I_{\text{э}}'^2 m R_{\text{экв}}}{2}, \quad (12-27)$$

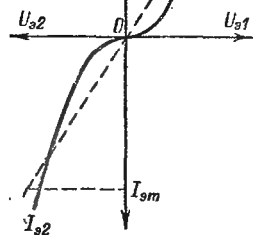


Рис. 12-11. Суммарная входная характеристика двухтактного каскада ОБ в классе В и ее идеализация для расчета входной мощности.

в которой сопротивление $R_{\text{экв}}$ больше, чем сопротивление $R_{\text{вх,ср}}$ на омическом участке (см. рис. 12-3).

Нелинейные искажения. В классе В даже при двухтактной схеме нелинейные искажения больше, чем в классе А. Коэффициент нелинейных искажений в значительной степени зависит от симметрии схемы, т. е. от равенства постоянных и переменных составляющих токов и напряжений в обоих плечах. Критерием симметрии обычно считают равенство постоянных составляющих коллекторных токов. Если эти составляющие различаются более чем на 20—30%, то необходимо осуществить искусственное выравнивание токов путем включения небольших сопротивлений последовательно с эмиттером того транзистора, у которого ток больше. Такой метод приводит к некоторому уменьшению к. п. д. и усиления, но позволяет заметно (в 2—3 раза) уменьшить нелинейные искажения и ток подмагничивания в выходном трансформаторе. При расчете симметрирующих сопротивлений можно пользоваться формулой (12-25) или (12-26).

Нелинейные искажения, возникающие во входной цепи, в значительной мере обусловлены изломом суммарной входной характе-

ристики вблизи нуля (рис. 12-11). Величину и влияние этого излома, особенно существенного у кремниевых транзисторов, можно значительно уменьшить, задавая небольшое положительное смещение на эмиттеры (рис. 12-12, а). Тогда суммарная характеристика будет более плавной, как видно из рис. 12-12, б.

Смещение E_3 лучше осуществлять не с помощью обычного делителя, а используя плоскостной диод в том плече, с которого снимается э. д. с. смещения, подобно тому, как это сделано на рис. 12-6. При наличии смещения каскад работает в классе АВ, поскольку ток покоя $I_{кА}$ может иметь значительную величину. При этом к. п. д. коллекторной цепи соответственно уменьшается, а ток и мощность в источнике питания возрастают. Эти изменения можно оценить

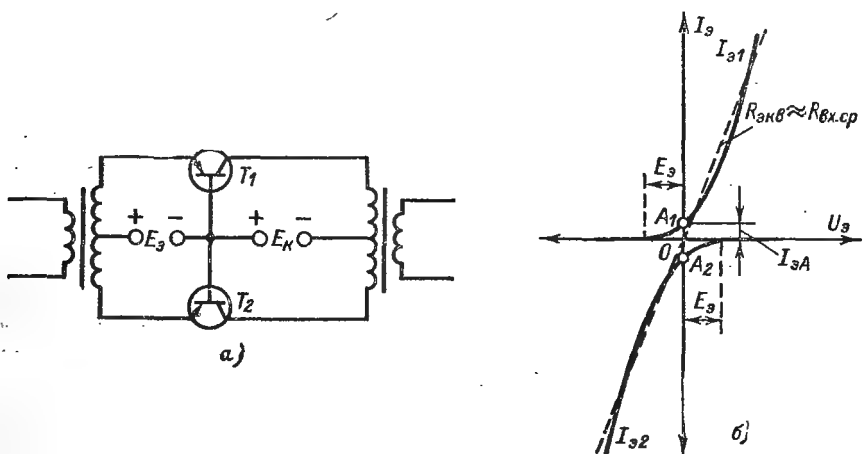


Рис. 12-12. Уменьшение входных нелинейных искажений с помощью смещения.
а — схема; б — суммарная входная характеристика.

по формулам (12-17) — (12-19), используя величину $I_{кА} \approx I_{к0} + \alpha I_{9А}$, где ток $I_{9А}$ легко найти из рис. 12-12, б. Входная мощность определяется формулой (12-27), в которой сопротивление $R_{экв}$ меньше, чем в схеме без смещения, и обычно близко к величине $R_{вх.ср}$ (см. рис. 12-3).

Общая особенность нелинейных искажений в двухтактных каскадах состоит в пониженной роли четных гармоник, прежде всего 2-й. В случае идеальной симметрии схемы четные гармоники вообще отсутствуют и расчет коэффициента нелинейных искажений по формуле (12-1) невозможен. В реальных, т. е. не совсем симметричных, схемах результирующая 2-я гармоника в нагрузке хотя и меньше, чем в каждом из плеч, но все же обычно существенна, так что коэффициент $K_{н.и}$ можно грубо оценить методом трех ординат (рис. 12-5), используя, конечно, составные характеристики.

Особенности двухтактного каскада, в котором транзисторы включены по схеме ОЭ, в основном те же, что и в одноконтном каскаде класса А (§ 12-1): меньшая входная мощность, несколько меньший к. п. д. и несколько бóльшие нелинейные искажения.

Следует отметить, что в двухтактных схемах ОЭ в отличие от одноконтных понятие оптимального сопротивления $R'_{г,опт}$ (см. конец § 12-1) имеет другой смысл: оно обеспечивает не равенство амплитуд положительной и отрицательной полувольт тока¹, а лишь меньшие искажения к а ж д о й из полувольт, т. е. уменьшение не ч е т н ы х гармоник. Обычно значение $R'_{г,опт}$ оказывается заметно бóльшим, чем $R_{экв}$, что невыгодно с точки зрения передачи мощности от предоконечного каскада. Поэтому, как правило, выбирают $R'_г \approx R_{экв}$, а коэффициент трансформации входного трансформатора находят из условия

$$n_1 = \frac{1}{2} \frac{\omega_2}{\omega_1} \approx \sqrt{\frac{R_{экв}}{R_г}}. \quad (12-28)$$

Это условие соответствует согласованию данного каскада с выходом предыдущего.

Важным средством уменьшения нелинейных искажений является использование отрицательной обратной связи — как общей, так и местной. Варианты схем мощных усилителей с местной обратной связью можно найти в [112].

Каскад с дополнительной симметрией. Выше рассматривались классические мощные усилители с трансформаторной связью, структура которых в сущности та же, что и при использовании ламп. Такое же структурное сходство свойственно и некоторым специальным типам мощных каскадов, применяемых в энергетических установках и автоматике [136]: усилителям среднего значения тока, усилителям с питанием от напряжения переменного тока и усилителям с импульсным режимом работы транзистора [137], в частности усилителям класса Д.

Однако в транзисторной технике возможны и некоторые оригинальные решения, к числу которых относятся двухтактные каскады, построенные на транзисторах с разным типом проводимости: $p-n-p$ и $n-p-n$. Сочетание транзисторов $p-n-p$ и $n-p-n$ используется не только в усилительной, но и в импульсной технике. Такие схемы получили название схем с *дополнительной симметрией*.

Двухтактный каскад с дополнительной симметрией показан на рис. 12-13. Характерная особенность каскада — отсутствие трансформаторов, что особенно важно в случае интегрального (микроэлектронного) исполнения, поскольку в интегральных схемах применение трансформаторов в виде навесных компонентов

¹ Выравнивание положительной и отрицательной амплитуд в двухтактных схемах достигается путем симметрирования.

крайне нежелательно, а их реализация в виде полупроводниковых или пленочных компонентов полностью исключается¹.

В отсутствие сигнала ток в нагрузке отсутствует, а через транзисторы протекает небольшой начальный ток, обусловленный смещениями $+\frac{1}{2}U_{R_2}$ (на транзисторе T_1) и $-\frac{1}{2}U_{R_2}$ (на транзисторе T_2)².

Эти смещения соответствуют э. д. с. E_s на рис. 12-12, б, т. е. каскад работает в классе АВ. Обычно ток делителя $R_1 - R_2 - R_3$ много больше начальных базовых токов, а сопротивление R_2 много меньше R_1 и R_3 . Поэтому начальные смещения $\pm \frac{1}{2}U_{R_2}$ можно считать равными $\frac{1}{2}IR_2$, где $I \approx 2E/(R_1 + R_3)$ — ток делителя.

Если сопротивление R_2 достаточно мало (а именно, если $R_2 \ll R_{вх1} \parallel R_{вх2}$, где $R_{вх}$ — входное сопротивление открытого плеча схемы), то можно считать обе базы соединенными по переменному току и подавать сигнал от предоконечного каскада на одну из баз (рис. 12-13)³. Поскольку входной сигнал приходит в фазе на оба транзистора, последние работают поочередно, как в обычном классе В (точнее, учитывая начальное смещение, — в классе АВ). При этом нагрузка включена в эмиттерные цепи транзисторов, т. е. они работают по схеме ОК, как эмиттерные повторители; следовательно, входной сигнал должен быть равен выходному напряжению на нагрузке. Соответственно, как отмечалось во введении, коэффициент усиления мощности оказывается меньше, чем при включении ОЭ.

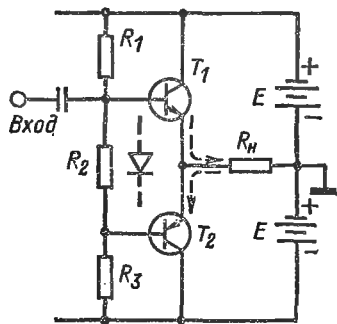


Рис. 12-13. Двухтактный каскад класса АВ с дополнительной симметрией.

Обычно вместо резистора R_2 ставят диод, работающий в прямом направлении (на рис. 12-13 показан пунктиром). В таком варианте начальные смещения на базах составляют $\pm U_d/2$, а роль сопротивления R_2 (при передаче сигнала на базу транзистора T_2) играет дифференциальное сопротивление диода. При достаточно больших токах дифференциальное сопротивление диода значи-

¹ Вообще говоря, используя транзисторы одного типа проводимости, можно построить схему так, чтобы избежать применения выходного трансформатора, но схема получается электрически несимметричной, а главное, возникают трудности, связанные с необходимостью противофазных входных сигналов, т. е. требуется либо входной трансформатор, либо специальный каскад для «расщепления» фазы.

² Средняя точка резистора R_2 имеет нулевой потенциал, что достигается регулировкой сопротивлений R_1 и R_3 . Если транзисторы идентичны, то $R_1 = R_3$.

³ В общем случае сигнал от предоконечного каскада следует подавать на обе базы через два отдельных конденсатора.

тельно меньше сопротивление постоянному току (см. рис. 2-22), поэтому использование диода позволяет усилить связь между базами транзисторов.

При расчете каскада с дополнительной симметрией роль сопротивления R'_H (рис. 12-9) играет сопротивление R_H . Поскольку отсут-

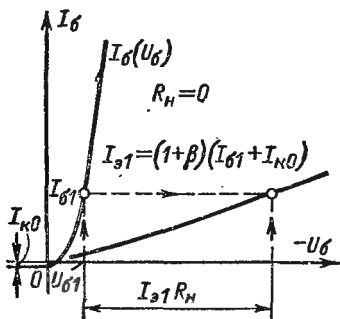


Рис. 12-14. Методика построения одной из точек «динамической» входной характеристики.

ствие трансформатора не позволяет искусственно увеличить сопротивление нагрузки, транзисторы работают с большими токами, чем в обычном трансформаторном каскаде. Однако еще более существенное различие касается входных характеристик: по оси абсцисс на рис. 12-11 и 12-12, б следует откладывать не напряжение $U_б$, а напряжение $U_б + I_б R_H$, поскольку потенциал эмиттера меняется вместе с потенциалом базы. Порядок перестройки входных характеристик показан на рис. 12-14. Перестроенные («динамические») характеристики позволяют оценить ме-

тодом пяти ординат [134] входную мощность и нелинейные искажения. В данном случае к. п. д. коллекторной цепи η_k совпадает с к. п. д. каскада в целом, поскольку отсутствуют потери в выходном трансформаторе и в цепи смещения.

Глава тринадцатая

УСИЛИТЕЛИ ПОСТОЯННОГО ТОКА

13-1. ВВЕДЕНИЕ

При разработке усилителей постоянного тока приходится решать две основные проблемы: согласование потенциалов в разных частях схемы и уменьшение нестабильности (дрейфа) выходного напряжения или тока в отсутствие сигнала.

Сначала остановимся на вопросах связи каскада с источником сигнала (рис. 13-1). Поскольку резистор $R_э$ является необходимым компонентом схемы [для стабилизации режима (см. гл. 6), а часто и для повышения входного сопротивления], то потенциал эмиттера и близкий к нему потенциал базы должны иметь определенное значение, отличное от нуля. С этой целью на рис. 13-1, а в цепь источника сигнала введена э. д. с. $E_б$ (ее можно осуществить с помощью делителя или полупроводникового стабилитрона, как показано пунктиром). Однако в такой схеме через источник сигнала будет протекать постоянная составляющая тока, — ток покоя базы, что часто нежелательно. Для того чтобы избежать тока через источник сигнала, можно задать потенциал базы $U_{б0}$ с помощью делителя

(рис. 13-1, б), а для компенсации тока через источник сигнала использовать э. д. с. $E_6 = U_{60}$. Если один из полюсов источника сигнала должен быть заземлен, то вход каскада можно выполнить по схеме, показанной на рис. 13-1, в где используется второй источник питания E'_K положительной полярности¹.

Тот же рис. 13-1 иллюстрирует способы согласования каскада с нагрузкой. На рис. 13-1, а и б в цепь нагрузки введена э. д. с. E_n , которая компенсирует коллекторный потенциал в режиме покоя (U_{KA}) и тем самым предотвращает протекание тока через нагрузку в отсутствие сигнала; э. д. с. E_n может быть получена теми же способами, что и E_6 . Во всех схемах заземление нагрузки возможно только при использовании «подвешенного» источника компенсирующего смещения (рис. 13-1, в).

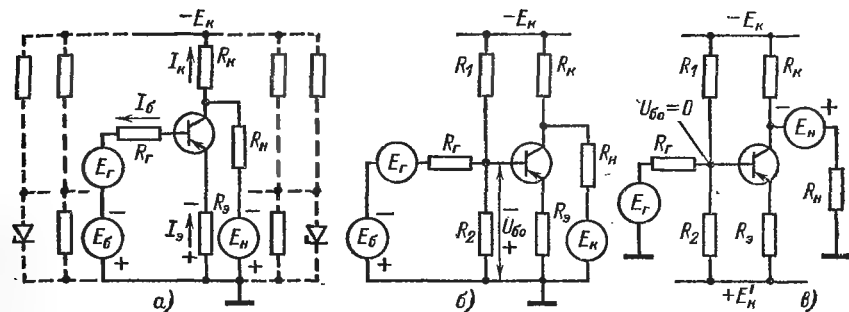


Рис. 13-1. Варианты согласования потенциальных уровней источника сигнала, нагрузки и каскада.

а — схема с заземленными источниками смещения; б — схема с заземленными источниками смещения и делителем в цепи базы; в — схема с двумя источниками питания и заземленным источником сигнала и нагрузкой.

В дальнейшем вопросы согласования каскада с источником сигнала и нагрузкой будем опускать, считая, что они решаются одним из описанных способов, и на схемах не будем показывать элементы E_r , R_r и R_n .

Проблема стабильности в транзисторных усилителях постоянного тока имеет свою специфику по сравнению с проблемой дрейфа в ламповых схемах. Это объясняется тем, что на положение рабочей точки помимо хаотических колебаний напряжения источников питания и хаотической «ползучести» параметров влияет еще и температура. Температурная зависимость характеристик и параметров, несмотря на индивидуальный количественный разброс у отдельных транзисторов, подчиняется определенным законам, которые были рассмотрены в гл. 2 и 4. Поэтому температурный дрейф в отличие от временной «ползучести» может быть в значительной степени уменьшен схемным путем. Прежде чем рассмотреть эти схемы, остановимся на количественной оценке температурного дрейфа.

¹ Значение э. д. с. E'_K и в целом параметры входной цепи рассчитывают из условия $U_{60} = 0$. Потенциал U_{60} можно записать, пользуясь теоремой узлового напряжения в форме (15-116), где под током I следует понимать ток покоя базы. Делитель R_1 — R_2 в таком варианте в сущности не нужен, так как база заземляется через источник сигнала.

Абсолютную величину дрейфа в транзисторном каскаде чаще всего выражают приращением коллекторного тока δI_K . В § 6-2 величина δI_K обозначилась через ΔI_K и была найдена как функция приращений ΔI_{K0} , $\Delta U_{эб}$ и $\Delta \beta$. Теперь учтем влияние еще одного параметра, зависящего от температуры, — сопротивления коллекторного перехода.

Как известно, сопротивление r_K влияет на коэффициент передачи базового тока. Это влияние можно проанализировать с помощью выражения (7-20). Дифференцируя (7-20) по r_K , нетрудно привести соответствующее частичное приращение к виду¹

$$(\Delta \beta_{oe})_{r_K} = \beta \gamma_K^* (1 - \gamma_K^*) \frac{\Delta r_K}{r_K}. \quad (13-1a)$$

Второе частичное приращение находим, дифференцируя (7-20) по β .

$$(\Delta \beta_{oe})_{\beta} = \Delta \beta \gamma_K^*. \quad (13-16)$$

Сложив выражения (13-1), найдем полное приращение $\Delta \beta_{oe}$ и, поделив его на β_{oe} , получим:

$$\frac{\Delta \beta_{oe}}{\beta_{oe}} = \frac{\Delta \beta}{\beta} + (1 - \gamma_K^*) \frac{\Delta r_K}{r_K}. \quad (13-2)$$

Еще одно влияние, которое оказывает сопротивление r_K на дрейф, связано с наличием составляющей коллекторного тока, обусловленной этим сопротивлением и источниками питания. Для анализа этого влияния воспользуемся схемой на рис. 13-2, а. Здесь для общности, как и на рис. 6-2, а, показаны э. д. с. во всех трех ветвях, но положено $U_{эб} = 0$ (тем самым считается, что R_6 — внешнее сопротивление в цепи базы). Преобразуем входную цепь по теореме об эквивалентном генераторе. Складывая э. д. с. E_K и э. д. с. эквивалентного генератора, получаем схему на рис. 13-2, б, где

$$E = E_K + \gamma_3 E_3 + \gamma_6 E_6;$$

коэффициенты γ_3 и γ_6 определяются обычными соотношениями. Применим метод суперпозиции. Положим сначала, что обратная связь отсутствует ($\beta = 0$); тогда

$$I'_K = \frac{E}{R_K + r_K^* + R_3 \parallel R_6}.$$

С учетом обратной связи ($\beta \neq 0$) ток I_K , как нетрудно показать, будет в $1 + \beta_{oe} \gamma_6$ раз меньше. Беря производную dI_K/dr_K и

¹ В усилителях постоянного тока нет блокирующих емкостей, поэтому коэффициент γ_K^* следует рассчитывать с учетом как внутренних, так и внешних сопротивлений схемы. А именно, в формулу (7-19a) вместо сопротивлений r_3 и $R_3 + r_6$ нужно подставить R_3 и R_6 (см. рис. 6-2).

полагая $\beta_{oe}\gamma_b \gg 1$, можно привести приращение $(\delta I_k)_E$ к следующему простому виду:

$$(\delta I_k)_E \approx -\frac{E}{r_k} \frac{\beta_{oe}}{1 + \beta_{oe}\gamma_b} \frac{\Delta r_k}{r_k}. \quad (13-3)$$

Подставив теперь в формулу (6-8) правую часть (13-2) вместо $\Delta\beta/\beta$ и добавив слагаемое (13-3), получим абсолютный температурный дрейф каскада:

$$\begin{aligned} \delta I_k = S \left\{ \Delta I_{k0} + \frac{\varepsilon \Delta T}{R_{\Sigma 6}} + \frac{I_{\Sigma}}{1 + \beta} \frac{\Delta\beta}{\beta} + \right. \\ \left. + \left[\frac{I_{\Sigma}}{1 + \beta} (1 - \gamma_k^*) - \frac{E}{r_k} \right] \frac{\Delta r_k}{r_k} \right\} = S \delta I_T, \end{aligned} \quad (13-4)$$

где положено $\Delta U_{\Sigma 6} = -\varepsilon \Delta T$.

Ток δI_T является по смыслу температурным дрейфовым током, соответствующим значению $S = 1$, а коэффициент нестабильности S

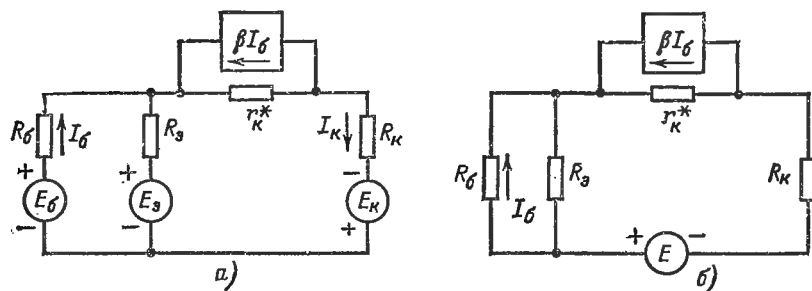


Рис. 13-2. Обобщенная эквивалентная схема каскада постоянного тока.

а — исходная схема; б — схема, преобразованная по теореме об эквивалентном генераторе.

определяется эквивалентным коэффициентом β_{oe} и является обобщением выражения (6-7):

$$S = \frac{\beta_{oe}}{1 + \beta_{oe}\gamma_b} = \frac{\beta\gamma_k^*}{1 + \beta\gamma_k^*\gamma_b}. \quad (13-5)$$

Обычно четвертое слагаемое в выражении (13-4) заметно меньше третьего, поэтому пренебрежение сопротивлением r_k в гл. 6 было оправдано. В усилителях же постоянного тока целесообразно учитывать все причины дрейфа. Из выражения (13-4) видно, что член с Δr_k может иметь разные знаки в зависимости от соотношения обоих слагаемых; в частности, он может быть равен нулю.

До сих пор мы рассматривали абсолютный дрейф. Однако на практике важнее знать относительный (приведенный) дрейф, т. е. дрейф, пересчитанный ко входу. Эту величину удобно сравнивать с величиной входного сигнала и, таким образом, оценивать чувствительность усилителя.

Покажем, что уменьшение коэффициента неустойчивости не решает проблемы относительного дрейфа.

В самом деле, определим из рис. 13-2 коэффициент усиления каскада, например, по базовому входу. Для этого используем формулу (7-16), положив $\beta = \beta_{oe}$, $R_n = \infty$ и $R_c + r_6 = R_6$; тогда

$$K_u = \frac{\Delta U_K}{\Delta E_6} = \frac{-\beta_{oe} R_K}{(R_3 + R_6)(1 + \beta_{oe} \gamma_6)} = -S \frac{R_K}{R_3 + R_6}. \quad (13-6)$$

Умножая (13-4) на R_K и деля на $|K_u|$, можно представить приведенный температурный дрейф в виде эквивалентного напряжения:

$$\delta U_{вх} = (R_3 + R_6) \delta I_T. \quad (13-7)$$

Как видим, приведенный дрейф действительно не зависит от коэффициента неустойчивости и уменьшается только с уменьшением сопротивлений R_3 и R_6 . Подставляя в (13-7) ток δI_T из (13-4) и полагая $R_3 + R_6 \rightarrow 0$, легко убедиться, что напряжение приведенного дрейфа стремится к пределу:

$$\delta U_{вх, \min} = |\Delta U_{эб}| = \epsilon \Delta T. \quad (13-8)$$

Например, если $\epsilon = 1,6$ мВ/град и $\Delta T = 100^\circ \text{C}$, то $\delta U_{вх, \min} = 160$ мВ. Следовательно, *температурная зависимость $U_{эб}(T)$ имеет принципиальное значение в усилителях постоянного тока, ограничивая их чувствительность.*

Согласно (13-8) низкоомные усилители на германиевых и на кремниевых транзисторах оказываются практически равноценными в отношении дрейфа¹: приведенный дрейф в диапазоне $\pm 60^\circ \text{C}$ составляет около 0,2 В или (при входном сопротивлении 50—100 кОм) несколько микроампер.

Необходимо подчеркнуть, что пропорциональность между коэффициентом усиления и коэффициентом неустойчивости, лежащая в основе выражения (13-7), свойственна только усилителям постоянного тока, поскольку в них отсутствуют блокирующие емкости. В усилителях переменного тока эти емкости позволяют обеспечить глубокую обратную связь по постоянной составляющей тока (т. е. малое значение S) и незначительную обратную связь по переменной составляющей тока (т. е. большое значение K_u).

13-3. ОДНОТАКТНЫЕ УСИЛИТЕЛИ

В настоящее время в усилителях постоянного тока, как правило, используются дифференциальные, т. е. двухтактные каскады (см. гл. 14). В данном параграфе мы рассматриваем одноконтурные усилители, главным образом для иллюстрации тех проблем, которые возникали при отказе от дифференциальных каскадов, что в прошлом, в дискретной (неинтегральной) схемотехнике, мотивировалось уменьшением числа транзисторов.

¹ Низкоомные сопротивления обеспечивают малую роль теплового тока — главного фактора температурной неустойчивости для германиевых транзисторов.

На рис. 13-3 показана схема трехкаскадного усилителя с непосредственной связью коллектора одного транзистора с базой другого. Такой способ каскадирования наиболее прост и экономичен, но требует соблюдения определенных условий. А именно, исходя из очевидных соотношений

$$|U_{кv}| = |U_{б(v+1)}| < |U_{к(v+1)}|;$$

$$|U_{эv}| < |U_{кv}| = |U_{б(v+1)}| \approx |U_{э(v+1)}|$$

и принимая, что токи у всех транзисторов одинаковы, нетрудно получить следующие неравенства¹:

$$R_{к1} > R_{к2} > R_{к3};$$

$$R_{э1} < R_{э2} < R_{э3}.$$

Эти неравенства означают, что коэффициент усиления от каскада к каскаду убывает: $K_{v+1} > K_v$ [см. (13-11)].

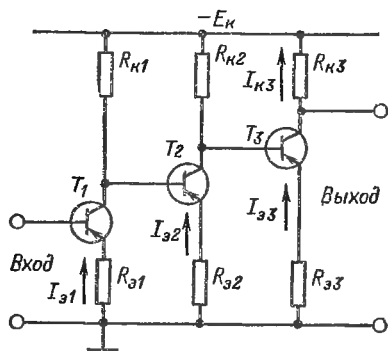


Рис. 13-3. Трехкаскадный однотактный усилитель постоянного тока.

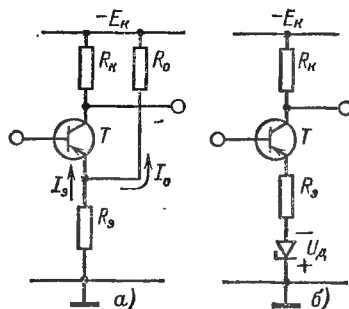


Рис. 13-4. Методы смещения эмиттеров в многокаскадном усилителе, начиная со второго каскада.

Если уменьшать ток в каждом следующем каскаде, то можно выровнять сопротивления $R_{к}$, но одновременно придется увеличить сопротивления $R_{э}$. Если, наоборот, увеличивать токи в каждом следующем каскаде, то можно выровнять $R_{э}$, но одновременно придется уменьшить сопротивления $R_{к}$. Поэтому (при заданном $E_{к}$) более рационально повышать потенциалы эмиттера от каскада к каскаду не путем увеличения сопротивления $R_{э}$, а с помощью дополнительных токов, поступающих через балластные сопротивления (рис. 13-4, а), или с помощью дополнительных напряжений, получаемых от полупроводниковых стабилитронов (рис. 13-4, б).

¹ В самом деле, $R_{кv} = (E_{к} - U_{кv})/I_{к}$ и $R_{эv} = U_{эv}/I_{э}$, где потенциалы $U_{кv}$ и $U_{эv}$ возрастают (по модулю) от каскада к каскаду.

Часто используется сочетание обоих способов, особенно тогда, когда сопротивление R_0 (рис. 13-4, а) получается сравнимым с R_3 (что снижает стабильность) или когда эмиттерный ток слишком мал для работы стабилитронов¹.

Добавление балластных сопротивлений и опорных напряжений не вносит принципиальных изменений в анализ переменных составляющих, т. е. дрейфа и усиления сигнала.

Напряжение питания E_k должно быть достаточно большим для установления нормального режима в последнем, самом «высоковольтном» каскаде:

$$E_k = I_{\beta 1} R_{\beta 1} + (U_{кз 1} + U_{кз 2} + U_{кз 3}) + I_{кз} R_{кз}. \quad (13-9)$$

Во всех каскадах, кроме первого, можно считать $R_{\beta v} = R_{к(v-1)}$. Поскольку R_{β} и $R_{к}$ имеют обычно сравнимые значения, коэффициент γ_6 редко бывает меньше 0,1—0,2. Поэтому, как правило, выполняется условие $\beta_{oe} \gamma_6 \gg 1$. Будучи использовано в формулах (13-5) и (13-6), оно приводит к следующим выражениям для коэффициентов неустойчивости и усиления каждого каскада:

$$S \approx \frac{1}{\gamma_6} = 1 + \frac{R_6}{R_{\beta}}; \quad (13-10)$$

$$K_u \approx \frac{R_{к}}{R_{\beta}} \quad (13-11)$$

В последнем для простоты опущен знак минус, ср. с (7-4, б)!. Максимальное усиление свойственно первому каскаду, у которого сопротивление $R_{к}$ имеет наибольшее значение. Однако даже в последнем каскаде, у которого $R_{к}$ минимально, должно соблюдаться условие $R_{к} > R_{\beta}$.

Рассмотрим результирующий дрейф в схеме на рис. 13-3. Легко заметить, что положительное приращение $\delta I_{к1}$ вызовет отрицательное приращение токов $I_{\beta 2}$ и $I_{к2}$. Таким образом, результирующий дрейф второго каскада может оказаться меньше, чем в отсутствие первого каскада, а при надлежащем подборе параметров он может быть сведен к нулю [71; 111]. В трехкаскадном усилителе такая возможность существенно затруднена, так как дрейф первого каскада усиливается больше, чем дрейф второго, и после усиления имеет ту же полярность, что и дрейф третьего каскада. Поэтому частичная или полная компенсация дрейфа свойственна, как правило, усилителям с четным числом каскадов; но и в этом случае она сопряжена с значительными трудностями (см. 1-е и 2-е издания данной книги).

¹ В качестве стабилитронов могут применяться и обычные кремниевые диоды, работающие в прямом направлении (если необходимое напряжение невелико и может быть реализовано с помощью небольшого числа диодов). В этом случае коэффициент ϵ в формуле (13-4) возрастает, тогда как в случае «настоящих» полупроводниковых стабилитронов, у которых $\epsilon_{ст} > 0$ (см. § 3-3), результирующий коэффициент ϵ обычно меняет знак и температурная стабильность при этом может ухудшиться.

Под термокомпенсацией понимают применение специальных температурно-зависимых схемных элементов, способствующих уменьшению температурного дрейфа. Такими элементами могут быть либо термисторы, либо полупроводниковые диоды. Последние имеют то преимущество, что температурные зависимости их обратного тока, прямого напряжения и соответствующих сопротивлений — такие же, как у транзисторов. Это облегчает компенсацию дрейфа в широком температурном диапазоне. Кроме того, диоды могут работать со значительно большими токами, чем термисторы, а это существенно в случае мощных усилителей.

Температурный коэффициент сопротивления у обоих отмеченных элементов отрицательный. Для термисторов он составляет около 3%/град. Для диодов лучше пользоваться зависимостями $I_{обр}(T)$ и $U_{пр}(T)$, описанными в § 2-6 и 2-7, не переходя к соответствующим сопротивлениям. При этом расчеты оказываются проще, чем при исполь-

зовании сопротивлений, температурные коэффициенты которых сильно зависят от токов и напряжений. Последовательно-параллельные комбинации диодов или термисторов с обычными линейными сопротивлениями позволяют получить оптимальную температурную зависимость в каждом конкретном случае [112].

Общий принцип термокомпенсации состоит в том, что диод или термистор включают в ту ветвь схемы, сопротивление которой желательно уменьшать с ростом температуры. Такую ветвь определяют путем качественного анализа схемы. Можно, например, включить термистор вместо R_K или в сочетании с ним. Тогда при надлежащем подборе параметров можно добиться условия

$$\Delta U_K = \Delta I_K R_K - I_K \Delta R_K = 0,$$

при котором выходное напряжение не будет меняться с изменением температуры. Однако чаще включают компенсирующий элемент во входную цепь транзистора. Примеры такой компенсации показаны на рис. 13-5. Здесь с ростом температуры уменьшается (по модулю) потенциал базы и соответственно уменьшается или полностью устраняется дрейфовое приращение тока I_K . С понижением температуры такой же эффект достигается за счет увеличения (по модулю) потенциала базы. Аналогичные решения применяются и в многокаскадных усилителях.

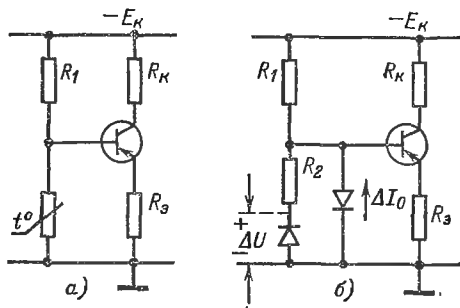


Рис. 13-5. Термостабилизация каскада с помощью термистора (а) и полупроводниковых диодов (б).

Все методы температурной стабилизации, описанные в предыдущих параграфах, подразумевают выполнение следующего принципиального условия: температура всех схемных элементов одинакова и меняется по одному и тому же закону.

В реальных усилителях это условие выполняется лишь с тем или иным приближением, поскольку имеет место саморазогрев элементов (особенно в мощных каскадах), существует градиент температуры внутри корпуса прибора, скаывается тепловая инерционность элементов и т. п. Учет всех факторов очень сложен и вряд ли оправдан на практике.

Сложной проблемой, не решаемой описанными усилителями, является временная дрейф. Правда, в дифференциальных усилителях (гл. 14) он меньше, чем в однотактных; однако сколько-нибудь надежная компенсация временного дрейфа, учитывая его нерегулярность и «индивидуальность», невозможна.

Еще одна проблема в области усилителей постоянного тока — низкочастотные шумы — также не решается описанными схемами. В самом деле, эти схемы характерны гальваническими, безреактивными связями, и следовательно, нулевой нижней

граничной частотой. Поэтому весь спектр низкочастотных шумов транзисторов (см. § 4-9) усиливается и беспрепятственно проходит на выход. Эти шумы нередко оказываются более существенными, чем температурный дрейф.

Отмеченные затруднения, связанные с температурным и временным дрейфом, а также с шумами, в значительной мере преодолеваются в усилителях постоянного тока с модуляцией входного сигнала. Блок-схема такого усилителя показана на рис. 13-6. Здесь сигнал постоянного тока сначала преобра-

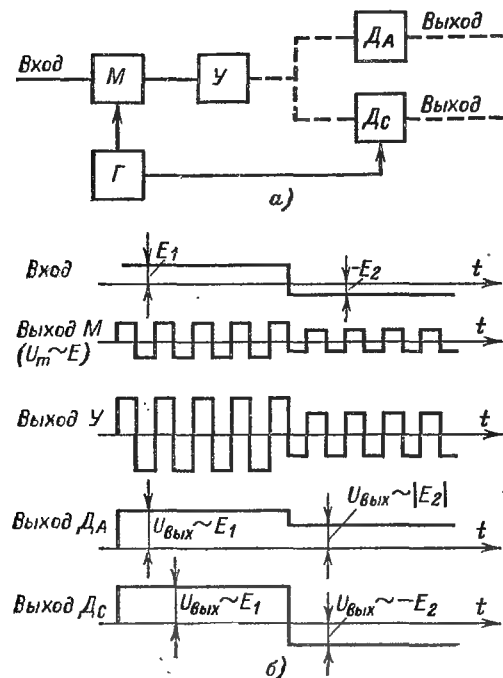


Рис. 13-6. Блок-схема усилителя с модуляцией входного сигнала (а) и ее временные диаграммы (б) (на рис. 13-6, а пунктиром показаны два варианта демодуляторов-детекторов: асинхронный $Д_A$ и синхронный $Д_C$).

зается в пропорциональный ему сигнал переменного тока (с помощью модулятора M), потом усиливается усилителем переменного тока $У$, а затем снова преобразуется в сигнал постоянного тока (с помощью демодулятора-детектора D). Поскольку в усилителях переменного тока с емкостной (или трансформаторной) связью дрейф рабочей точки не передается от каскада к каскаду, основной проблемой при данном методе усиления является только стабильность, «бездрейфовость» самого модулятора. Нижняя граничная частота усилителя $У$ отлична от нуля, а потому спектр низкочастотных шумов ограничен и их роль уменьшается.

Усилители постоянного тока с промежуточной модуляцией (M) входного сигнала и последующей демодуляцией (DM) усиленного выходного сигнала называют *усилителями типа М-ДМ*. В отличие от них «обычные» усилители постоянного тока (со сквозными гальваническими связями) называют *усилителями прямого усиления*.

Общий недостаток усилителей типа М-ДМ заключается в малом значении верхней граничной частоты (не более 0,1—0,2 частоты модуляции), а также в необходимости иметь спе-

циальный источник модулирующего напряжения (G). Однако в интегральных микросхемах последний недостаток несуществен, так как площадь, занимаемая указанным источником на кристалле, составляет лишь малую долю всей площади схемы.

В качестве модуляторов могут использоваться любые элементы, параметрами которых (сопротивлением, емкостью или индуктивностью) можно управлять с помощью вспомогательного источника переменного тока.

Контактные прерыватели типа поляризованного реле (рис. 13-7, а) обеспечивают максимальную «глубину модуляции» (так как сопротивление цепи меняется от нуля до бесконечности). Однако они имеют существенные недостатки: ограниченную рабочую частоту (обычно 50—100 Гц), недостаточную надежность (срок службы) и значительные помехи—наводки от катушки электромагнита (требуется громоздкая экранировка).

Все три отмеченных недостатка практически отсутствуют при использовании транзисторных прерывателей (рис. 13-7, б). В этой схеме происходят поочередное отпирание (до насыщения) и запи- рание транзисторов под действием модулирующего напряжения, т. е. они действуют как ключи в цепи модулируемого сигнала. Ключевой режим транзисторов подробно рассматривается в гл. 15, а сами прерыватели — в § 15-3. Сейчас мы только отметим, что

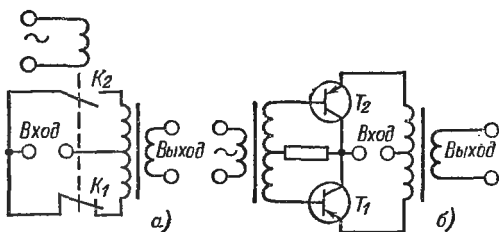


Рис. 13-7. Модуляторы-прерыватели.

а — контактный; б — транзисторный.

глубина модуляции в случае транзисторных прерывателей несколько меньше, чем в случае механических контактов, особенно при малом сигнале. Кроме того, транзисторным прерывателям свойствен большой временной и особенно температурный дрейф. Зато большая рабочая частота транзисторных прерывателей (сотни и тысячи герц) позволяет делать достаточно высокой и нижнюю граничную частоту усилителя U и тем самым не пропускать основную часть спектра низкочастотных шумов транзистора. Одновременно повышается верхняя граничная частота усиливаемых сигналов.

В последнее время большое внимание уделяется прерывателям на МДП транзисторах [100, 138]. Такие прерыватели, как отмечалось в § 5-5, характеризуются отсутствием остаточного напряжения в «замкнутом» состоянии и поэтому обеспечивают большую глубину модуляции и меньшие значения модулируемого сигнала (см. § 15-3).

Помимо транзисторных прерывателей можно отметить такие полупроводниковые модуляторы, как емкость p - n перехода [139, 140] и датчик Холла [141]. Для работы последнего необходимо создавать магнитное поле. Что касается диодного модулятора, то он основан на наличии у p - n перехода нелинейной зависимости $C_n(U)$ согласно (2-82). Правда, эта зависимость слаба и не дает возможности получить большую глубину модуляции. Однако следует учитывать наличие в настоящее время других типов полупроводниковых нелинейных емкостей, отличных от p - n переходов и имеющих более крутую вольт-фарадную характеристику [142].

Выходной демодулятор-детектор может быть асинхронным и синхронным. В первом случае это обычный двухполупериодный или мостовой выпрямитель с фильтром. Во втором случае это такой же прерыватель, как на входе, питаемый той же частотой. Синхронный детектор в отличие от асинхронного позволяет различать полярность входного сигнала постоянного тока (см. рис. 13-6, б).

Одной из главных проблем при разработке М-ДМ усилителей является борьба с импульсами, проходящими от генератора G (рис. 13-6, а) через паразитные емкости модулятора на вход усилителя U и далее, после усиления, на выход. Эти импульсы имеют вид коротких «пиков» чередующейся полярности, которые соответствуют положительному и отрицательному фронтам прямоугольных импульсов, вырабатываемых генератором. При неизбежном различии амплитуд «пиков» появляется разностная постоянная составляющая, «маскирующая» полезный сигнал. Наиболее эффективным способом борьбы с паразитными «пиками» является использование в модуляторе приборов с минимальными паразитными емкостями. В последнее время в качестве таких приборов выступают МДП транзисторы с самосовмещенным затвором (с. 313).

ДИФФЕРЕНЦИАЛЬНЫЙ КАСКАД

14-1. ВВЕДЕНИЕ

Простейшая схема дифференциального каскада (ДК) показана на рис. 14-1. Название «дифференциальный» объясняется тем, что выходное напряжение (между коллекторами) пропорционально разности входных потенциалов (на базах). Дифференциальные каскады известны давно, еще в эпоху ламповой техники, где их называли параллельно-балансными каскадами. Однако наибольшее распространение они приобрели в конце 60-х годов в связи с развитием интегральных операционных усилителей. Несмотря на то, что по своей структуре ДК являются классическими представителями усилителей постоянного тока, они благодаря ряду уникальных свойств используются в гораздо более широком аспекте, а в микроэлектронике стали одним из универсальных элементов линейных интегральных схем.

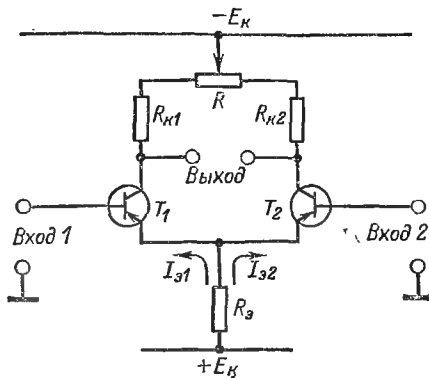


Рис. 14-1. Дифференциальный каскад постоянного тока.

Операционным усилителям и их основе — ДК посвящена обширная литература (см. [139, 143] и библиографию в них), поэтому ниже даются лишь самые необходимые сведения.

14-2. ОБЩИЕ СВОЙСТВА

Из рис. 14-1 видно, что ДК симметричен. Именно симметрия позволяет существенно уменьшить величину приведенного дрейфа по сравнению с однотактным каскадом. Вторым важным преимуществом ДК является значительно более высокий коэффициент усиления. Рассмотрим подробнее эти две особенности.

Если параметры обоих плеч ДК полностью идентичны (включая их зависимость от температуры и времени), то, очевидно, выходной дрейф будет строго равен нулю. В реальных схемах полная идентичность, разумеется, исключена и имеется некоторый «разностный» дрейф. Теоретически этот дрейф рассчитать невозможно, но известны пути его уменьшения или, что то же самое, пути искусственного симметрирования ДК.

Действительно, условием отсутствия дрейфа является равенство $\delta I_{к1} R_{к1} \approx \delta I_{к2} R_{к2}$. Подставляя сюда $\delta I_{к} = S \delta I_T$ согласно (13-4)

и полагая $S \approx 1/\gamma_6$ (что обычно выполняется), запишем условие нулевого дрейфа в виде ¹

$$\frac{R_{k1}}{R_{k2}} = \frac{\delta I_{T1}}{\delta I_{T2}} \frac{2R_3 + R_{62}}{2R_3 + R_{61}} \approx \frac{\delta I_{T1}}{\delta I_{T2}}. \quad (14-1)$$

В выражении (14-1) значения δI_T — заданные; по существу заданы также значения R_6 . Сопротивление R_3 менять нежелательно, так как оно оказывает влияние на другие важные параметры (см. ниже); кроме того, обычно выполняется неравенство $2R_3 \gg R_6$, что использовано при упрощении правой части (14-1). Следовательно, основным средством компенсации дрейфа является регулировка соотношения сопротивлений R_k , например, с помощью потенциометра R (рис. 14-1).

Нужно иметь в виду, что помимо условия нулевого дрейфа должно выполняться условие нулевого выходного напряжения в отсутствие сигнала: $I_{k1}R_{k1} = I_{k2}R_{k2}$. Вообще говоря, последнее равенство не коррелирует с (14-1), т. е. в схеме ДК необходима еще одна «степень свободы» (еще одна регулировка). Обычно потенциометр R используется для регулировки нуля напряжения, а для регулировки симметрии используются дополнительные цепи, на которых мы не останавливаемся ².

Симметрия ДК существенно улучшается, а диапазон входных сигналов увеличивается, если между эмиттерами включить небольшое сопротивление — «связку» (рис. 14-2). Наличие

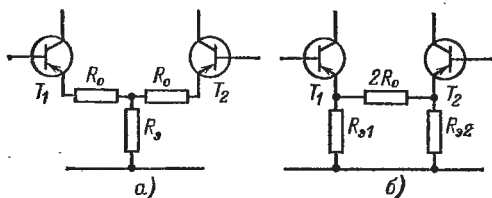


Рис. 14-2. Варианты стабилизации дифференциального каскада по переменному току.

а — с последовательной связкой; б — с параллельной связкой.

такой «связки» [144] равносильно искусственному увеличению сопротивлений r_3 на величину R_0 ; если $R_0 \gg r_3$, то сопротивления эмиттеров делаются линейными и стабильными, а возможное различие в значениях r_{31} и r_{32} сглаживается. Увеличение диапазона входных сигналов

объясняется тем, что напряжение на э к в и в а л е н т н о м эмиттерном переходе (включающем резистор R_0) больше, чем на истинном переходе, на величину $I_3 R_0$. Поэтому, если без «связки» (см. § 10-3) диапазон входных сигналов не превышает $2U_{эб}$, то при нали-

¹ При расчете коэффициентов γ_{61} и γ_{62} следует мысленно разделить R_3 на два параллельно включенных сопротивления $R_{31} = R_{32} = 2R_3$ и отнести каждое из них к соответствующей половине схемы. Что касается сопротивлений R_6 , то они включают в себя сопротивления баз и сопротивления источников сигнала. Если источник сигнала включен между базами, то его сопротивление следует поделить поровну между обоими входами.

² В интегральных ДК внутренние регулировки всякого рода исключены, зато асимметрия плеч оказывается существенно меньше, чем в дискретной схеме. Поэтому регулировка дрейфа (симметрии) отсутствует, а балансировка нуля осуществляется в н е ш н е й схемой.

чий «связки» он расширяется до $2(U_{\text{зб}} + I_{\text{з}}R_0)$. Обычно сопротивление R_0 составляют несколько сотен ом (при токах около 1 мА).

В схеме на рис. 14-2, а резисторы R_0 можно выполнить в виде потенциометра с движком посередине; тогда его можно одновременно использовать для регулировки дрейфа или «нуля» наряду или вместо потенциометра R на рис. 14-1.

Заметим, что иногда выходное напряжение в ДК снимают не между коллекторами, а между эмиттерами. Тогда сопротивление нагрузки оказывается включенным так, как резистор $2R_0$ на рис. 14-2, б, т. е. оно одновременно выполняет и функцию «связки». Тем самым обеспечиваются высокая стабильность схемы и широкий динамический диапазон. Эти свойства объясняются тем, что в данном случае ДК представляет собой балансный (двухтактный) э м и т т е р - н ы й п о в т о р и т е л ь (сопротивления R_k в такой схеме отсутствуют).

Перейдем к оценке коэффициента усиления ДК в основной схеме (рис. 14-1). Пусть входной сигнал $U_{\text{вх}}$ поступает между базами. Тогда на одном из входов (относительно «земли») будет действовать напряжение $+U_{\text{вх}}/2$, а на другом $-U_{\text{вх}}/2$. Соответственно $\Delta I_{\text{э1}} = -\Delta I_{\text{э2}}$, т. е. *сумма эмиттерных токов не изменяется, а значит, не изменяется и потенциал эмиттеров*. Отсюда следует, что обратная связь через резистор $R_{\text{э}}$ отсутствует и коэффициент усиления каждого плеча будет таким же, как в о б ы ч н о м каскаде ОЭ (без резистора $R_{\text{э}}$). Его можно рассчитать по формуле (7-4а), если под сопротивлением нагрузки понимать $R_{\text{н}}/2$ (поскольку потенциал средней точки нагрузки не меняется). Очевидно, что *коэффициент усиления у ДК намного больше, чем у однотактного каскада*, где он ограничен величиной $R_{\text{э}} \gg r_{\text{э}}$ [см. (13-11)].

Анализ показывает (см. ниже), что если входной сигнал распределяется между плечами не поровну или (в пределе) поступает только на одну из баз, то сделанные выводы относительно отсутствия обратной связи через резистор $R_{\text{э}}$ и соответственно идентичности коэффициентов усиления у ДК и у обычного каскада ОЭ остаются в силе.

ДК можно соединять друг с другом н е п о с р е д с т в е н н о: коллекторы предыдущего каскада с базами следующего (рис. 14-5). При этом, поскольку сопротивления $R_{\text{э}}$ не оказывают влияния на коэффициент усиления, можно повышать эмиттерные потенциалы каждого следующего каскада простейшим способом — увеличением сопротивлений $R_{\text{э}}$ без применения дополнительных элементов, свойственных однотактным схемам (см. рис. 13-4). Это еще одно существенное преимущество ДК.

Как правило, 1-й (входной) ДК в многокаскадном усилителе рассчитывают так, чтобы он не нуждался в цепях смещения баз, т. е. *начальные базовые потенциалы 1-го ДК равны нулю*. Это позволяет работать от заземленных источников сигнала (см. рис. 13-1, в и сноску на с. 429). Из аналогичных соображений последний (выходной) ДК делают с н е с и м м е т р и ч н ы м выходом, т. е. снимают выходное напряжение между одним из коллекторов и землей. Это позволяет использовать наиболее распространенную з а з е м л е н н у ю нагрузку. Нередко роль такой нагрузки играет некоторый промежуточный однотактный каскад (см. § 14-6).

При анализе ДК удобно и общепринято разделять входные и выходные сигналы на две составляющие: *дифференциальную* и *синфазную*. Последнюю называют также *средним уровнем* или просто *уровнем*¹.

Эти две составляющие выражаются через приращения входных и выходных потенциалов с помощью следующих определений:

$$U_{\text{вх. у}} = \frac{1}{2} (\Delta U_{61} + \Delta U_{62}); \quad U_{\text{вх. д}} = \Delta U_{61} - \Delta U_{62}; \quad (14-2a)$$

$$U_{\text{вых. у}} = \frac{1}{2} (\Delta U_{к1} + \Delta U_{к2}); \quad U_{\text{вых. д}} = \Delta U_{к1} - \Delta U_{к2}. \quad (14-2б)$$

Обратные соотношения имеют вид:

$$\Delta U_{61} = U_{\text{вх. у}} + \frac{1}{2} U_{\text{вх. д}}; \quad \Delta U_{62} = U_{\text{вх. у}} - \frac{1}{2} U_{\text{вх. д}}; \quad (14-3a)$$

$$\Delta U_{к1} = U_{\text{вых. у}} + \frac{1}{2} U_{\text{вых. д}}; \quad \Delta U_{к2} = U_{\text{вых. у}} - \frac{1}{2} U_{\text{вых. д}}. \quad (14-3б)$$

Связь между выходными и входными сигналами ДК выражается системой двух уравнений:

$$U_{\text{вых. у}} \approx K_{\text{у. у}} U_{\text{вх. у}} + K_{\text{у. д}} U_{\text{вх. д}}; \quad (14-4a)$$

$$U_{\text{вых. д}} = K_{\text{д. у}} U_{\text{вх. у}} + K_{\text{д. д}} U_{\text{вх. д}}, \quad (14-4б)$$

где $K_{\text{у.у}}, \dots, K_{\text{д.д}}$ — коэффициенты передачи соответствующих составляющих со входа на выход.

В идеальном (абсолютно симметричном) ДК соблюдается условие $K_{\text{у.д}} = K_{\text{д.у}} = 0$, т. е. *передачи синфазной и дифференциальной составляющих осуществляются независимо друг от друга*. Собственно в этом и состоит удобство разделения сигналов на указанные составляющие. Рассмотрим параметры идеального ДК.

1) Если подать на входы ДК одинаковые по величине, но разные по полярности напряжения, то согласно (14-2a) синфазная составляющая на входе будет отсутствовать: $U_{\text{вх.у}} = 0$. Поскольку $K_{\text{у.д}} = 0$, то будет отсутствовать и синфазная составляющая на выходе [см. (14-4a)]. Следовательно, в данном случае на выходе будет только усиленная дифференциальная составляющая входного сигнала:

$$U_{\text{вых. д}} = K_{\text{д. д}} U_{\text{вх. д}}. \quad (14-5)$$

Дифференциальный коэффициент усиления $K_{\text{д.д}}$ (обычно индексы опускают) является одним из основных параметров ДК. Есте-

¹ Термин «уровень» менее распространен, но индекс «у» более удобен, чем «с» (синфазный), который можно спутать, например, с обозначением для «стока». Поэтому ниже для синфазных составляющих используется индекс «у», а для дифференциальных — «д».

венно, его стараются делать как можно больше. Выше было показано, что усиление каждого плеча ДК такое же, как в обычном каскаде ОЭ. Значит, коэффициент $K_{д.д}$ можно записать, например, в форме (7-20а)*:

$$K_{д.д} = \frac{-\beta_{oe}}{1 + \beta_{oe}\gamma_6} \frac{R_K \parallel \left(\frac{1}{2} R_H\right)}{R_r + r_6 + r_3}, \quad (14-6a)$$

где использование величины $\frac{1}{2} R_H$ обосновано в предыдущем разделе.

При наличии «связки» (см. рис. 14-2) к сопротивлению r_3 добавляется сопротивление R_0 , что, конечно, приводит к снижению коэффициента усиления. Полагая в этом случае $\beta_{oe}\gamma_6 \gg 1$, получаем:

$$K_{д.д} \approx - \frac{R_K \parallel \left(\frac{1}{2} R_H\right)}{R_0 + r_3}. \quad (14-6б)$$

Помимо улучшения симметрии ДК (см. предыдущий раздел), наличие сопротивлений R_0 приводит к существенному росту входного сопротивления:

$$R_{вх.д} = r_6 + \beta_{oe}(R_0 + r_3); \quad (14-7)$$

это часто является дополнительным преимуществом, искупающим спад усиления.

2) Если соединить входные зажимы ДК и подать на них общее напряжение, то согласно (14-2а) дифференциальная составляющая на входе будет отсутствовать: $U_{вх.д} = 0$. Поскольку $K_{д.у} = 0$, то будет отсутствовать и дифференциальная составляющая на выходе [см. (14-4б)]. Следовательно, в данном случае на выходе будет только та или иная часть синфазной составляющей входного сигнала (усиление обычно не имеет места):

$$U_{вых.у} \approx K_{у.у} U_{вх.у}. \quad (14-8)$$

Сдвиг выходного у р о в н я в виде напряжения $U_{вых.у}$ и изменение коллекторных токов (одинаковое в обоих плечах) означают изменение рабочего режима транзисторов (сдвиг рабочей точки). Поэтому коэффициент передачи $K_{у.у}$ желательно делать как можно меньше.

В рассматриваемом случае базы и эмиттеры обоих транзисторов соединены друг с другом; коллекторы также можно считать соединенными (так как разность потенциалов $U_{к1} - U_{к2}$ сохраняет нулевое значение в силу симметрии). Значит, в режиме синфазного сигнала ДК представляет собой параллельное соединение обоих плеч. Расчет такой схемы несложен: он аналогичен расчету каскада с обратной связью по току (рис. 8-5). Используя выражение (8-22б)

* Поскольку на вход каждого плеча подается половина сигнала $U_{вх.д}$, а на выходе усиленные сигналы складываются (т. е. удваиваются), то коэффициент усиления ДК в целом равен коэффициенту усиления каждого плеча.

(поскольку нагрузка R_n в данном режиме «бездействует»), получаем синфазный коэффициент усиления:

$$K_{y,y} \approx -\frac{R_k}{2R_s}. \quad (14-9)$$

Здесь коэффициент 2 обусловлен параллельным соединением коллекторных резисторов, а сопротивления r_s и R_0 не учтены как малые по сравнению с R_s . Обычно $R_k < R_s$, так что $K_{y,y} < 1$.

Входное сопротивление для синфазной составляющей получается из формулы (8-19):

$$R_{вх,y} = \frac{1}{2} r_6 + \beta_{oe} \left(R_s + \frac{1}{2} r_s \right) \approx \beta_{oe} R_s. \quad (14-10)$$

Очевидно, что $R_{вх,y} \gg R_{вх,д}$ [см. (14-7)] из-за существенного различия в значениях R_s и R_0 .

До сих пор рассматривался идеальный ДК. В реальных ДК условие $K_{y,д} = K_{д,y} = 0$ не выполняется из-за неизбежной асимметрии схемы. Конечные, хотя и весьма малые, значения взаимных коэффициентов передачи означают, что дифференциальный сигнал $U_{вх,д}$ несколько меняет выходной уровень [см. (14-4а)], а синфазная составляющая $U_{вх,y}$ вносит некоторый вклад в дифференциальный выходной сигнал [см. (14-4б)]. Первое обстоятельство мало существенно, зато второе представляет серьезную проблему, так как во многих случаях имеет место соотношение $U_{вх,y} \gg U_{вх,д}$ и тогда «паразитная» составляющая $K_{д,y} U_{вх,y}$ может оказаться больше полезной составляющей $K_{д,д} U_{вх,д}$.

Параметр $K_{д,y}$ как таковой практически не используется; его принято характеризовать косвенно, с помощью так называемого коэффициента подавления (или режекции) синфазной составляющей (англ. CMRR — Common Mode Rejection Ratio):

$$K_{\pi} = |K_{д,д} / K_{д,y}|.$$

Коэффициент подавления характеризует относительную роль слагаемых в выражении (14-4б) при равенстве дифференциального и синфазного входных сигналов¹. Как правило, его оценивают в логарифмических единицах — децибелах:

$$K_{\pi} = 20 \lg \left| \frac{K_{д,д}}{K_{д,y}} \right|. \quad (14-11)$$

Например, если $|K_{д,д} / K_{д,y}| = 10\,000$, то $K_{\pi} = 80$ дБ.

Распространенным типом синфазного сигнала являются различные (внутренние и внешние) помехи и наводки, действующие одновременно на оба входа. Поэтому вопрос увеличения коэффициента K_{π} есть одновременно вопрос повышения помехоустойчивости ДК.

¹ Исходя из этого, нетрудно показать, что обратная величина коэффициента подавления характеризует различие в усилении асимметричного сигнала (поданного на один из входов) и симметричного (поданного между обоими входами).

Поскольку конечное значение K_n обусловлено несимметрией ДК, а асимметрия в свою очередь обусловлена очень многими (иногда трудно учитываемыми) факторами, то теоретический расчет коэффициента K_n практически невозможен. В то же время из общих соображений можно сделать некоторые **качественные** выводы. Например, любые меры, заведомо способствующие повышению симметрии ДК, приводят к росту коэффициента K_n . Это подтверждается, в частности, тем, что в интегральных ДК коэффициент подавления заметно выше, чем в дискретных (если, конечно, исключить подбор транзисторов). Далее, поскольку степень симметрии, в силу зависимостей $\beta(I_k, T)$, $U_{ab}(I_k, T)$ и др., меняется с изменением температуры и рабочего режима ДК, то и коэффициент подавления должен зависеть от этих факторов. Практика подтверждает, что значение K_n довольно существенно зависит от рабочих токов и напряжений, причем в ряде случаев удается выявить **оптимальный режим**. Подтверждается также зависимость K_n от температуры.

Заметим, что целью анализа коэффициента K_n является не столько возможность его численного расчета¹, сколько выявление тех путей, которые позволяют увеличить K_n при заданной степени асимметрии ДК. Опуская довольно громоздкие выкладки (см. [139]), приведем результат анализа:

$$K_n = \frac{\beta_{oe} R_3 / |\delta|}{R_r + R_{вх.д}} \approx \frac{1}{|\delta|} \frac{R_3}{R_0 + r_3 + (R_r + r_6) / \beta_{oe}}. \quad (14-12)$$

Здесь параметр δ представляет собой алгебраическую сумму относительных разбросов величин β , R_k , R_r и др. Значение δ следует считать заданной.

Главный вывод, который следует из выражения (14-12), состоит в том, что *токозадающий резистор R_3 должен иметь как можно большее сопротивление*. Роль сопротивлений R_r и r_6 незначительна, а сопротивление R_0 не должно быть излишне большим.

В заключение заметим, что в случае несимметричного выхода (например, если выходной сигнал снимается с одного из коллекторов, см. конец предыдущего параграфа) степень симметрии теряет свое значение и коэффициент подавления определяется совсем иначе:

$$K_n^* \approx \frac{\frac{1}{2} K_{д.д}}{K_{у.у}}.$$

Здесь коэффициент $\frac{1}{2}$ обусловлен тем, что каждое из плеч ДК усиливает **половину** входного дифференциального сигнала. Подставляя значения (14-6), (14-9) и полагая для простоты $R_n = \infty$, получаем после преобразований²:

$$K_n^* = \frac{R_3}{R_0 + r_3 + \frac{R_r + r_6}{\beta_{oe}}}. \quad (14-13)$$

¹ На практике этот коэффициент всегда измеряется экспериментально.

² Если **учитывать нагрузку**, то следует в формуле (14-6) заменить $R_n/2$ на R_n (так как при несимметричном выходе нагрузка **полностью** подключена параллельно R_k), а в формуле (14-9) заменить $R_k/2$ на $(R_k/2) \parallel R_n$.

Сравнивая (14-13) и (14-12), видим, что оба коэффициента подавления связаны соотношением $K_n^* = |\delta|K_n$, т. е. $K_n^* \ll K_n$.

В области высших частот и малых времен ДК не имеет специфики по сравнению с каскадом ОЭ. Частотные и переходные характеристики дифференциального коэффициента усиления могут рассчитываться по формулам, изложенным в гл. 7, а характеристики синфазного коэффициента усиления — по тем же формулам, но с использованием постоянной времени (8-23).

14-4. ТОЧНОСТНЫЕ ПАРАМЕТРЫ

Под точностными параметрами будем понимать такие, которые определяют погрешности ДК, специфичные для него как для усилителя постоянного тока. К числу таких параметров относятся: *начальный разбаланс входного напряжения (или напряжение смещения нуля)* и его температурный дрейф, *средний входной ток и разбаланс входного тока*. Рассмотрим кратко эти параметры.

Начальным разбалансом напряжения U_p называют величину входного напряжения, при которой выходное напряжение равно нулю (имеются в виду дифференциальные составляющие обоих напряжений). Поскольку в идеальном (симметричном) ДК нулевому выходному напряжению соответствует нулевое входное, можно считать, что напряжение U_p , как и коэффициент $K_{д.у.}$, характеризует асимметрию ДК. Однако, в отличие от $K_{д.у.}$, параметр U_p характеризует асимметрию не по переменным, а по постоянным составляющим токов и напряжений (или как говорят, по постоянному току).

Начальный разбаланс всегда определяется экспериментально; так как, будучи связан с асимметрией схемы, он зависит от многих трудно учитываемых факторов и не поддается точному расчету. Однако полезно провести следующий упрощенный анализ, который хорошо иллюстрирует суть проблемы и одновременно дает полезные результаты.

Пусть коллекторные токи и сопротивления в обоих плечах ДК идентичны; тогда $I_{к1}R_{к1} = I_{к2}R_{к2}$ и $U_{вых} = 0$. Если бы напряжения на обоих эмиттерных переходах тоже были идентичны, то на вход не нужно было бы подавать никакого корректирующего напряжения, чтобы сохранить условие $U_{вых} = 0$. Тогда начальный разбаланс U_p был бы равен нулю. На деле имеет место некоторый разброс напряжений $U_{эб}$ при одинаковых токах $I_э$ из-за различия тепловых токов [см. (4-9)]. Чтобы скомпенсировать указанный разброс, необходимо подать на вход напряжение

$$U_p = U_{эб1} - U_{эб2}, \quad (14-14)$$

которое по определению и будет начальным разбалансом. Используя выражение (4-9) и полагая для простоты $I'_{э0} = I_{э0}$, получаем:

$$U_p = \varphi_T \ln \frac{I_{э02}}{I_{э01}}. \quad (14-15)$$

Например, если тепловые токи различаются на 20%, то напряжение разбаланса составит $U_p \approx 5$ мВ — значение, типичное для многих ДК. В интегральных схемах, где транзисторы расположены рядом друг с другом на одной подложке, вероятность различия электрофизических параметров полупроводника

на смежных участках очень невелика; поэтому различие тепловых токов, а вместе с тем и начальный разбаланс будут меньше, чем в дискретных схемах¹.

Дифференцируя (14-14) по температуре, получаем температурную чувствительность начального разбаланса. В правой части оказывается разность температурных чувствительностей, для которых можно использовать выражение (2-66); тогда

$$\epsilon_{U_p} = \frac{U_p}{T}. \quad (14-16)$$

Например, если $U_p = 5$ мВ и $T = 300^\circ\text{C}$, то $\epsilon_{U_p} \approx 17$ мкВ/°C.

Выражение (14-16) отражает важный факт: *уменьшение начального разбаланса сопровождается уменьшением его температурного дрейфа*. Значит, в интегральных схемах дрейф меньше, чем в дискретных. Однако выражение (14-16) теряет силу при значениях $U_p < 1$ мВ, так как в исходном выражении (2-66) не учтены некоторые факторы, которые несущественны по отношению к величине $U_{\text{об}}$, но становятся существенными по отношению к малой разности $\Delta U_{\text{об}} = U_p$.

Если необходимо уменьшить начальный разбаланс по сравнению с его «естественным» значением, применяют в *внешней цепи* балансировки нуля.

Входными токами ДК являются базовые токи транзисторов. Для простоты положим $I_{k0} = 0$, т. е. будем рассматривать кремниевые транзисторы. Тогда $I_b = I_k/\beta$ и средний входной ток можно записать в следующем виде:

$$I_{\text{вх. ср}} = \frac{1}{2} \left(\frac{I_{k1}}{\beta_1} + \frac{I_{k2}}{\beta_2} \right) \approx \frac{I_k}{\beta} = \frac{I_0}{2\beta}, \quad (14-17)$$

где I_k — усредненный коллекторный ток, а β — усредненный коэффициент усиления. Такое усреднение допустимо потому, что разброс параметров в плечах ДК невелик.

Роль среднего входного тока как параметра ДК обусловлена тем, что он, протекая по сопротивлению источника сигнала, дает падение напряжения, равносильное синфазной составляющей. Следовательно, при достаточно больших значениях $I_{\text{вх. ср}}$ и R_g может, во-первых, появиться паразитный дифференциальный сигнал на выходе и, во-вторых, измениться рабочий режим каскада (см. предыдущий параграф). Таким образом, ток $I_{\text{вх. ср}}$ ограничивает допустимое значение R_g , поэтому его следует уменьшать.

Что касается *разбала* входных токов (или разностного входного тока)

$$I_p = \frac{I_{k1}}{\beta_1} - \frac{I_{k2}}{\beta_2}, \quad (14-18)$$

то он полностью зависит от разброса токов и коэффициентов усиления. Если принять, что коллекторные токи выровнены тем или

¹ Главным фактором начального разбаланса в интегральных ДК становится различие геометрических размеров (площадей переходов), обусловленное погрешностями фотолитографии.

иным путем, разбаланс входных токов можно записать следующим образом:

$$I_p = \frac{\Delta\beta}{\beta_1\beta_2} I_K \approx \frac{\Delta\beta}{\beta} I_{\text{вх. ср}}, \quad (14-19)$$

где β — усредненный коэффициент усиления.

Из выражения (14-19) видно, что разбаланс токов меньше среднего тока (так как $\Delta\beta \ll \beta$) и уменьшается вместе с последним. Видно также, что важную роль играет идентичность коэффициентов усиления β . Это лишний раз подтверждает рациональность интегрального исполнения ДК.

Роль разбаланса среднего тока как параметра состоит в том, что при протекании по сопротивлению источника сигнала ток I_p дает падение напряжения, равносильное дифференциальному сигналу или начальному разбалансу входного напряжения. Поэтому уменьшение тока I_p является одной из первоочередных задач.

Считая, что температурные зависимости среднего тока и его разбаланса сосредоточены в коэффициенте β , нетрудно, дифференцируя (14-17) и (14-19), получить соответствующие температурные чувствительности:

$$\epsilon_{I_{\text{вх. ср}}} = \left(\frac{1}{\beta} \frac{d\beta}{dT} \right) I_{\text{вх. ср}} = B I_{\text{вх. ср}}; \quad (14-20a)$$

$$\epsilon_{I_p} \approx \left(\frac{1}{\beta} \frac{d\beta}{dT} \right) I_p = B I_p. \quad (14-20б)$$

Обе эти температурные чувствительности, как и ϵ_{U_p} , пропорциональны самим параметрам. Что касается коэффициента B , то он согласно [139] меняется от $-1,5$ до $-1,0\%/^{\circ}\text{C}$ в диапазоне от -60 до 0°C и от $-1,0$ до $-0,5\%/^{\circ}\text{C}$ в диапазоне от 0 до $+125^{\circ}\text{C}$.

Помимо рассмотренных усилительных и точностных параметров ДК характеризуются еще многими другими (выходные сопротивления, коэффициент подавления изменений питающих напряжений, максимальные входные и выходные сигналы и т. д.). На них мы не будем останавливаться. Заметим только, что максимально допустимые входные сигналы (как дифференциальный, так и синфазный) ограничены в сущности нелинейными искажениями, а в пределе — отсечкой или насыщением транзисторов.

14-5. ЭВОЛЮЦИЯ СХЕМЫ И ПАРАМЕТРОВ

Как уже отмечалось, прогресс ДК в последнее время связан с развитием микроэлектроники. Последняя, в частности, характерна коренным пересмотром понятия «сложности схемы»: в отличие от дискретной схемотехники критичным стало количество пассивных компонентов (конденсаторов и резисторов), а не активных — транзисторов и диодов. Поэтому в интегральных ДК транзисторы широко используются как в качестве диодов и резисторов,

Увеличение коэффициента подавления требует прежде всего увеличения токозадающего сопротивления R_3 [см. (14-12) и (14-13)]. В этом направлении возникают те же трудности, что и при увеличении коллекторных сопротивлений, поэтому и решение проблемы по аналогии состоит в использовании динамических нагрузок. В данном случае их называют источниками тока. На рис. 14-3 роль такого источника тока играет транзистор T_5 . Его выходное сопротивление можно оценить по формуле (7-23) с учетом резистора R_3 в эмиттерной цепи. Обычно сопротивление $R_{\text{вых}}$ близко к величине r_k , которая в микрорежиме согласно (4-24) может иметь значение до 10 МОм и более. Задаваемый ток можно оценить из соотношения

$$I_0 \approx \frac{E_6 - U_{эб5}}{R_3},$$

где для кремниевых транзисторов $U_{эб} \approx 0,7$ В.

Напряжение разбаланса и его температурный коэффициент, как и в простейшей схеме, уменьшают регулировкой коллекторных и эмиттерных резисторов (см. предыдущие параграфы).

Что касается входного тока и его разбаланса, то их уменьшение достигается как путем уменьшения коллекторных токов (микрорежим), так и путем увеличения коэффициента β (использование составных и супербета-транзисторов, см. § 4-10 и с. 240). На рис. 14-3 иллюстрируется использование составных транзисторов — пар Дарлингтона, каждая из которых обозначена как единый прибор. Согласно (14-20) вместе с уменьшением рассматриваемых параметров уменьшаются и их температурные коэффициенты.

В последние годы во входных ДК многокаскадных усилителей все шире используются полевые транзисторы (унитроны), которые обеспечивают очень малые входные токи (до 1 нА и менее), малые токи разбаланса (до единиц пикоампер), а также весьма высокие входные сопротивления (сотни мегаом). Однако при этом нерешенной проблемой остается сравнительно большое напряжение разбаланса, особенно в интегральных схемах, где исключен подбор транзисторов. Кроме того, полевые транзисторы, как известно (см. гл. 5), не обеспечивают таких высоких коэффициентов усиления, как биполярные.

В заключение приведем некоторые типичные параметры с о б р е м е н н ы х (многотранзисторных) ДК:

Напряжение питания E_K	12 В
Ток коллектора I_K	20—50 мкА
Коэффициент усиления $K_{д, д}$	50—100
Коэффициент подавления $K_{п}$	80—100 дБ
Коэффициент подавления $K_{п}^*$	60—70 дБ
Входное сопротивление $R_{вх. д}$	0,5—1 МОм
Входное сопротивление $R_{вх. у}$	10—100 МОм
Напряжение разбаланса U_p	1—5 мВ
Дрейф напряжения разбаланса U_p	3—10 мВ/град
Средний входной ток $I_{вх. ср}$	10—100 нА
Ток разбаланса I_p	0,5—5 нА

Для сравнения укажем, что простейший ДК (рис. 14-1) характеризуется параметрами: $K_{д.д} = 20 \div 50$; $K_{п} = 50 \div 60$ дБ, $R_{вх.д} = 10 \div 20$ кОм, $I_{вх.ср} = 5 \div 20$ мкА, $I_p = 0,5 \div 1$ мкА. Как видим, прогресс в отношении коэффициента подавления и входных величин весьма значителен.

14-6. ОПЕРАЦИОННЫЕ УСИЛИТЕЛИ

Дифференциальные каскады часто используют в измерительных схемах, когда нужно измерить разность потенциалов между двумя «незаземленными» точками. Однако наиболее широкое применение ДК нашли в операционных усилителях, особенно интегральных.

Операционным усилителем (ОУ) называют усилитель постоянного тока с дифференциальным входом и одноктактным выходом (рис. 14-5), имеющий весьма высокие коэффициент усиления и входное сопротивление.

Наличие дифференциального входа не означает необходимости работать только с дифференциальными сигналами, симметричными относительно земли. Один из входных зажимов можно заземлить, подавая сигнал на второй зажим от заземленного источника. В таком режиме различают *инвертирующий* и *неинвертирующий* входы, в зависимости от того, находится ли входной сигнал в противофазе или в фазе с выходным (рис. 14-4).

Типовым является использование ОУ в режиме параллельной отрицательной обратной связи по напряжению. На рис. 14-4 элементы цепи обратной связи — резисторы R_1 и R_2 , причем R_2 включает в себя сопротивление источника сигнала.

Из очевидных соотношений

$$E_r = U_{вх} + I_2 R_2; \quad U_{вых} = -I_1 R_1 + U_{вх}; \quad I_{вх} = \frac{U_{вх}}{R_{вх}} = I_2 - I_1$$

легко получить коэффициент усиления ОУ с обратной связью:

$$K_{о.с} = \frac{-R_1/R_2}{1 + \frac{1}{K} \left(1 + \frac{R_1}{R_2} + \frac{R_1}{R_{вх}} \right)}, \quad (14-21a)$$

где $K = |U_{вых}/U_{вх}|$ — коэффициент усиления ОУ без обратной связи.

Из полученного выражения ясно видна причина, по которой стремятся увеличивать параметры K и $R_{вх}$: при этом уменьшается

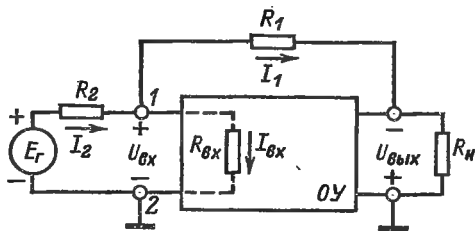


Рис. 14-4. Блок-схема операционного усилителя.

1 — инвертирующий вход; 2 — неинвертирующий вход.

погрешность (в виде 2-го члена в знаменателе), связанная с влиянием температуры, режима и т. п.

В идеальном ОУ, у которого $R_{вх} = \infty$ и $K = \infty$, имеем:

$$K_{о.с} = -\frac{R_1}{R_2}, \quad (14-216)$$

т. е. коэффициент усиления зависит только от параметров внешней цепи обратной связи.

Помимо усиления ОУ может выполнять многие другие функции. Так, если заменить резистор R_1 на конденсатор C , то коэффициент усиления будет выражаться операторной функцией интегрирования ($1/sCR_2$); такой ОУ называют *интегрирующим усилителем*. Если заменить конденсатором C резистор R_2 , получим функцию дифференцирования (sCR_1) и ОУ будет *дифференцирующим усилителем*. При условии $R_1 = R_2$ получаем *инвертирующий повторитель*. При подаче комплексной обратной связи на неинвертирующий вход схема легко самовозбуждается и превращается в тот или иной тип генератора. Именно такая гибкость ОУ сделала его основным, универсальным компонентом линейных (аналоговых) интегральных схем.

Типовая структура ОУ включает в себя один или, чаще, два ДК ($T_1 - T_4$ на рис. 14-5), один переходный одноктактный каскад (T_5) и выходной двухтактный каскад (T_6, T_7). Схема последнего рассматривалась в § 12-3.

Операционный усилитель описывается той же системой параметров, что и ДК, причем свойства ОУ прежде всего определяются свойствами входного ДК. Поэтому в последнем используются все меры, способствующие улучшению входных параметров (см. § 14-5). Второй и последующие каскады призваны в основном обеспечить нужный коэффициент усиления и выходные параметры (напряжение, ток, выходное сопротивление). Современные ОУ характеризуются коэффициентами усиления $10^4 - 10^5$ и выше, выходными токами 10—20 мА и выходными напряжениями 5—10 В (последние, конечно, зависят от напряжения питания).

Рис. 14-5. Внутренняя структура операционного усилителя.

14-7. КАСКАД НА МДП ТРАНЗИСТОРАХ

В основе ДК — его обоих плеч — лежат, как известно, одноктактные каскады. Поэтому начнем с их рассмотрения

Простейшие одноктактные каскады на МДП транзисторах могут иметь три разновидности. У первой из них (рис. 14-6, а) в стоковой части используется обычный резистор, у остальных двух — динамическая нагрузка в виде МДП тран-

зистора с тем же (рис. 14-6, б) или противоположным (рис. 14-6, в) типом проводимости канала.

Каскад с резисторной нагрузкой мало специфичен по сравнению с классическим ламповым каскадом (если, конечно, не учитывать температурные зависимости, шумы и другие эксплуатационные факторы). Поэтому к нему можно применить всю хорошо разработанную теорию ламповых схем (см. например, [62]). На этом каскаде мы не будем останавливаться, ограничившись выражением для коэффициента усиления:

$$K = -\frac{\mu}{1 + R_I/R_c}, \quad (14-22)$$

где μ — коэффициент усиления транзистора; R_I — его внутреннее сопротивление [см. (5-61) и петит на с. 336]. При достаточно большом сопротивлении нагрузки ($R_c \gg R_I$) получаем максимальный коэффициент усиления каскада: $K \approx \mu$. Обычно $K = (0,2 \div 0,4) \mu$.

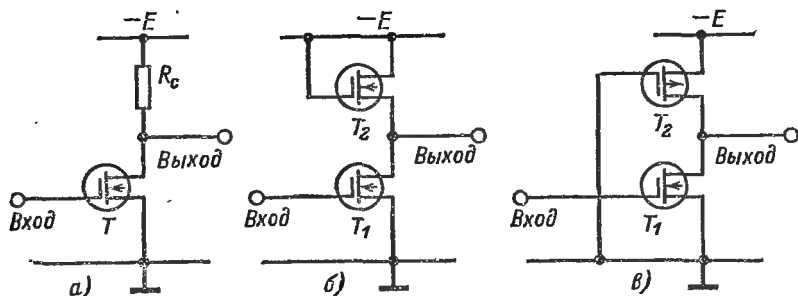


Рис. 14-6. Простейшие однотактные каскады на МДП транзисторах.

а — с резисторной нагрузкой; б — с транзисторной нагрузкой того же типа проводимости; в — с транзисторной нагрузкой противоположного типа проводимости.

Рассмотрим каскад на однотипных транзисторах (рис. 14-6, б). Из условия $U_{с2} = U_{с2}$ следует: $U_{с2} > U_{з2} - U_{02}$, т. е. [см. (5-39)] транзистор T_2 работает в области насыщения, на пологих участках характеристик (рис. 5-28, а) *. Соответственно для обоих транзисторов будем пользоваться формулами насыщенного режима; для простоты ограничимся 1-м приближением, т. е. формулами (5-60) — (5-63). Как и в гл. 5, во избежание знаков минуса будем записывать потенциалы и напряжения в виде модулей.

Роль сопротивления R_c в рассматриваемой схеме играет входное сопротивление каскада T_2 , включенного по схеме ОЗ. Следовательно, в формуле (14-22) можно заменить R_c на $R_{вхОЗ}$. Используя упрощенное выражение (5-90) и подставляя $R_{д1} = \mu_1/S_1$, получаем:

$$K = -\frac{\mu_1}{\mu_1 \frac{S_2}{S_1} + 1}. \quad (14-23)$$

Отношение крутизн запишем с помощью выражения (5-63). Поскольку токи обоих транзисторов одинаковы, легко находим:

$$\frac{S_2}{S_1} = \sqrt{\frac{b_2}{b_1}} = \frac{1}{\sqrt{B}}. \quad (14-24)$$

* Задавая независимое смещение на затвор T_2 , можно обеспечить ненасыщенный режим, но при этом параметры каскада ухудшаются (см. [145]). Кроме того, недостатком будет наличие дополнительного источника смещения.

Здесь отношение удельных кривизн обозначено специальным символом

$$B = \frac{b_1}{b_2}. \quad (14-25)$$

Если в выражении (5-55) считать величины ϵ , μ (подвижность) и d одинаковыми для обоих транзисторов (что обычно выполняется), то коэффициент B определяется только геометрическими факторами — соотношением длин и ширины каналов.

Таким образом, в общем случае

$$B = \frac{Z_1}{Z_2} \frac{L_2}{L_1}, \quad (14-26a)$$

а в частном случае, при $L_1 = L_2$,

$$B = \frac{Z_1}{Z_2}. \quad (14-26b)$$

Вариант (14-26b) особенно ясно иллюстрирует, что значения коэффициента B непосредственно связаны с «размерами» транзисторов; если $B > 1$, то транзистор T_1 «больше», чем T_2 ; если $B < 1$, то транзистор T_1 «меньше» транзистора T_2 . В гл. 5 отмечалось, что при линейной структуре транзисторов геометрический фактор Z/L ограничен: его трудно сделать более 100 и менее 1; значения B для такой структуры не превышают 40—50.

Подставляя (14-24) в (14-23), получаем:

$$K = \frac{-\mu_1}{\frac{\mu_1}{\sqrt{B}} + 1} \approx -\sqrt{B}. \quad (14-27)$$

Приближение основано на том, что обычно $\mu = 50 \div 100$ и более, а $\sqrt{B}$ не превышает 6-7.

Итак, коэффициент усиления в безрезисторной схеме на однотипных транзисторах очень невелик и принципиально ограничен геометрическими факторами. Для того чтобы получить $K \approx \mu_1$, необходимы значения $B \geq 10^4$, что нереально даже в гребенчатых структурах (см. сноску на с. 313).

Относительно статического режима каскада можно сделать следующие замечания. Запишем очевидное равенство

$$E = U_{c1} + U_{z2}, \quad (14-28)$$

где оба напряжения отсчитываются от истоков. Выражая токи обоих транзисторов с помощью (5-60), приравнявая их друг другу и подставляя коэффициент B из (14-25), получаем соотношение

$$\frac{U_{z2} - U_0}{U_{z1} - U_0} = \sqrt{B}. \quad (14-29)$$

Здесь пороговые напряжения обоих транзисторов приняты одинаковыми, что характерно для дискретных схем, в которых подложки можно соединять с истоками.

Подставляя U_{z2} из (14-29) в (14-28), приходим к выводу, что напряжение питания и напряжения U_{z1} , U_{c1} , определяющие рабочую точку активного транзистора T_1 , связаны однозначно:

$$E = \sqrt{B} U_{z1} + U_{c1} - (\sqrt{B} - 1) U_0. \quad (14-30a)$$

Если желательно выполнить условие $U_{z1} = U_{c1}$, то соотношение (14-30a) переходит в следующее¹:

$$E = (\sqrt{B} + 1) U_{z1} - (\sqrt{B} - 1) U_0. \quad (14-30b)$$

¹ Равенство $U_z = U_c$, если выполнить его во всех каскадах многокаскадного усилителя, позволяет сделать все каскады идентичными и соединять их друг с другом непосредственно без всякого согласования. При этом начальный выходной потенциал (в отсутствие сигнала) будет равен выход-

Учитывая, что величина $\sqrt{B}$ согласно (14-27) есть коэффициент усиления, можно сказать, что выражения (14-30) устанавливают связь между усилением и режимом каскада. Например, если принять $\sqrt{B} = |K| = 5$; $U_0 = 3$ В и $E = 10$ В, то из (14-30б) получим $U_{a1} = 4,25$ В; потенциал U_{c1} будет иметь такое же значение. Если при тех же начальных условиях нужно обеспечить $U_{c1} = 6$ В, то из (14-30а) получаем $U_{a1} = 3,45$ В.

Теперь перейдем к каскаду с разнотипными транзисторами (рис. 14-6, в). В этом случае нагрузочный транзистор T_2 включен тоже по схеме ОЗ, но роль сопротивления R_c в формуле (14-22) теперь играет не входное, а выходное сопротивление $R_{\text{выхОЗ}}$.

Если бы транзистор T_2 работал в насыщенном режиме, то имело бы место равенство $R_{\text{выхОЗ}} = R_{i2}$. Такой вариант (когда последовательно включены два весьма больших дифференциальных сопротивления, в данном случае R_{i1} и R_{i2}) характерен низкой стабильностью.

В ненасыщенном режиме¹ транзистора T_2 сопротивление $R_{\text{выхОЗ}}$ можно найти из выражения (5-69а); запишем его следующим образом:

$$R_{i2} = \frac{1}{b_2 (U_{a2} - U_0) (1 - v_2)}, \quad (14-31)$$

где $v_2 = \frac{U_{c2}}{U_{a2} - U_0}$ — коэффициент ненасыщенности ($0 < v_2 < 1$).

Приравняв токи I_{c1} и I_{c2} , из которых первый выражается формулой (5-60), а второй — формулой (5-58), можно получить соотношение, аналогичное (14-29):

$$\frac{U_{a2} - U_0}{U_{a1} - U_0} = \frac{\sqrt{B}}{\sqrt{v_2} (2 - v_2)}. \quad (14-32)$$

Подставляя (14-31) в (14-22) вместо R_c , заменяя R_{i1} на μ_1/S_1 , выражая S_1 согласно (5-62) и учитывая (14-32), приведем коэффициент усиления к виду

$$K = \frac{-\mu_1}{\frac{\mu_1 (1 - v_2)}{\sqrt{B v_2} (2 - v_2)} + 1} \approx -\sqrt{B} \frac{\sqrt{v_2} (2 - v_2)}{1 - v_2}. \quad (14-33)$$

Приближенная формула получена из условия $\sqrt{B} \ll \mu_1$, которое позволяет пренебречь единицей в знаменателе точной формулы при всех реальных значениях v_2 [реальные значения v_2 не превышают 0,8, в противном случае выражение (5-69а) дает большую погрешность].

Сравнивая (14-33) и (14-27), убеждаемся, что данная схема может иметь преимущество перед предыдущей: если, например, $v_2 = 0,7$, то $K \approx 3,2 \sqrt{B}$. С уменьшением коэффициента v_2 выигрыш уменьшается. Несмотря на это преимущество, схема с разнотипными транзисторами реже используется на практике в связи с тем, что качественные МДП транзисторы с n -каналом получены только в самое последнее время. Поэтому ниже в основу ДК положен каскад с однотипными p -канальными транзисторами.

н о м у. В частности, используя в цепях истоков дополнительный источник питания $+E_1$, можно получить н у л е в ы е (относительно земли) начальные потенциалы и на входе, и на выходе.

В случае биполярных транзисторов аналогичное равенство $U_6 = U_k$ невозможно, так как оно соответствует границе активного и насыщенного режимов. В активном режиме, когда $U_k > U_6$, возникают те проблемы согласования каскадов и их «индивидуализации» по смещениям, о которых говорилось в § 13-3.

¹ Условие ненасыщенного режима $U_{a2} - U_0 > U_{c2}$ в данной схеме обеспечивается при выполнении неравенства $U_{c1} > U_0$.

Простейший безрезисторный ДК на МДП транзисторах показан на рис. 14-7. Плечами этого ДК являются рассмотренные выше однотактные каскады (рис. 14-6, б). Транзистор T_0 , работающий с постоянным смещением, выполняет функцию источника тока, подобно транзистору T_5 на рис. 14-3. Цепь смещения транзистора T_0 для простоты не показана.

Анализ схемы с учетом разброса основных параметров плеч приводит к следующим выражениям для коэффициентов передачи [146]:

$$K_{y. y} = \frac{K_{cp}}{1 + 2S_{cp}R_{i0}}; \quad (14-34a)$$

$$K_{y. л} \approx \frac{1}{4} |K_{cp}| \delta K + \delta S; \quad (14-34б)$$

$$K_{л. y} = \frac{|K_{cp}|}{1 + 2S_{cp}R_{i0}} \delta K; \quad (14-34в)$$

$$K_{л. л} \approx K_{cp}. \quad (14-34г)$$

Рис. 14-7. Дифференциальный каскад на МДП транзисторах.

В формулах (14-34) приняты следующие обозначения. Для средних величин:

$$K_{cp} = \frac{1}{2} (K_1 + K_2); \quad S_{cp} = \frac{1}{2} (S_1 + S_2).$$

Для разбросов (допусков) относительно средних значений:

$$\delta K = \frac{K_1 - K_2}{K_{cp}}; \quad \delta S = \frac{S_1 - S_2}{S_{cp}}.$$

Индексами «1» и «2» обозначены величины, относящиеся соответственно к плечам, содержащим транзисторы T_1 и T_2 ; R_{i0} — внутреннее сопротивление транзистора T_0 . Под крутизнами S_1 и S_2 , строго говоря, понимаются не крутизны самих транзисторов, а крутизны однотактных каскадов, составляющих плечи ДК. Однако в данной схеме эта замена не существенна¹.

Деля (14-34г) на (14-34в), получаем коэффициент подавления синфазного сигнала:

$$K_{\pi} = \frac{1 + 2S_{cp}R_{i0}}{|\delta K|}. \quad (14-35)$$

Как видим, коэффициент подавления определяется относительным разбросом коэффициентов усиления плеч и произведением $S_{cp}R_{i0}$.

¹ Связь между крутизной самого транзистора S и крутизной каскада S^* имеет вид:

$$S^* = S \frac{R_i}{R_i + R_c}.$$

При типичном условии $R_c \ll R_i$, равносильном условию $K \ll \mu$ [см. (14-22)], получаем $S^* \approx S$.

Разброс δK зависит главным образом от разброса геометрических размеров — длин L и ширины Z каналов соответствующих транзисторов [см. (14-27) и (14-26)]. Что касается произведения $S_{\text{ср}} R_{i0}$, то возможности его увеличения будут яснее после некоторых преобразований.

А именно, положим, что оба плеча ДК идентичны и имеют средние параметры ($I_{\text{ср}}$, $S_{\text{ср}}$, $b_{\text{ср}}$). Тогда, записывая (5-63) для $S_{\text{ср}}$ и для S_0 и учитывая, что $I_0 = 2I_{\text{ср}}$, нетрудно получить соотношение

$$S_{\text{ср}} = \frac{S_0}{2\sqrt{B_0}}, \quad (14-36)$$

где

$$B_0 = \frac{1}{2} \frac{b_0}{b_{\text{ср}}}. \quad (14-37)$$

Теперь, подставляя (14-36) в (14-35) и пренебрегая единицей в числителе, можно записать коэффициент подавления в весьма наглядной форме:

$$K_{\text{п}} \approx \frac{\mu_0}{|\delta K| \sqrt{B_0}}. \quad (14-38)$$

Например, если $B_0 = 0,4$, $\mu_0 = 50$ и $\delta K = 0,02$, то $K_{\text{п}} \approx 3500$ (т. е. 70 дБ).

Очевидно, что для величины $K_{\text{п}}$ определяющую роль играет разброс коэффициентов усиления плеч. Учитывая (14-27), (14-25) и (5-55), можно сказать, что в общем случае этот разброс обусловлен допусками на подвижность, толщину диэлектрика и геометрический фактор Z/L . При интегральном исполнении ДК главное значение имеет последний допуск, при использовании дискретных транзисторов, изготовленных по одному и тому же фотолитографическому процессу, — первые два допуска, а при использовании произвольных транзисторов — все три допуска. Как правило, интегральное исполнение обеспечивает снижение разброса на порядок по сравнению с дискретным ДК на неподавленных транзисторах.

Из выражения (14-38) ясна роль еще двух параметров — μ_0 и B_0 . У большинства МДП транзисторов коэффициент μ имеет максимум при токах, в несколько раз меньших номинального (см. сноску на с. 306). Значит, если это допускают условия стабильности (с учетом малой разности $U_{\text{з}} - U_0$), следует ставить транзистор T_0 в режим пониженных токов.

Что касается коэффициента B_0 , то его уменьшение, диктуемое выражением (14-38), может быть достигнуто двумя путями: уменьшением удельной крутизны b_0 и увеличением удельной крутизны $b_{\text{ср}}$. Первый путь связан с уменьшением геометрического фактора Z_0/L_0 . Если длина канала L_0 такая же, как у основных транзисторов (5-10 мкм), то получить отношение $Z_0/L_0 < 1$ затруднительно по технологическим причинам (разрешающая способность фотолитографии). Если же увеличить длину канала L_0 , то может заметно

измениться конфигурация поля истоков-затвор, расположенная в основу всего анализа МДП транзисторов; кроме того, увеличивается «растекание» дырок из канала по оси z (см. рис. 5-29). Поэтому реальные значения Z_0/L_0 составляют не менее 0,2—0,3.

Следовательно, основным способом уменьшения коэффициента B_0 (ради увеличения K_n) является увеличение удельной крутизны b_{cp} активных транзисторов T_1 и T_2 . Это одновременно способствует повышению коэффициента усиления. Последний согласно (14-34г) и (14-27) можно записать в виде

$$K = - \frac{\mu_{cp}}{\frac{\mu_{cp}}{\sqrt{B_{cp}}} + 1} \approx - \sqrt{B_{cp}}, \quad (14-39)$$

где

$$B_{cp} = (b_1 + b_2)/(b_3 + b_4).$$

Как в однотактном каскаде [см. (14-30)], статический режим ДК оказывается связанным с параметром $\sqrt{B}$, т. е. с коэффициентом усиления; кроме того, на режим ДК влияет параметр $\sqrt{B_0}$. Опуская сравнительно громоздкие выкладки [146], приведем основные результаты анализа.

Если положить в начальном состоянии схемы $U_{31} = U_{32} = 0$ (что характерно для ДК), то соотношение между напряжениями питания имеет вид:

$$\frac{E_1 - U_0}{E_2 - U_0 - \Delta_0} = \frac{\sqrt{B_{cp}}}{1 + 1/\sqrt{B_0}}, \quad (14-40)$$

а необходимое напряжение смещения для транзистора T_0 определяется формулой¹

$$U_{30} = - \left(\frac{E_1 - U_0}{\sqrt{B_{cp}}} + \Delta_0 \right). \quad (14-41)$$

Например, если $\sqrt{B_{cp}} = |K_{cp}| = 5$; $\sqrt{B_0} = 0,5$; $U_0 = 3$ В и $\Delta_0 = 5$ В то из приведенных формул следует:

$$E_1 = 1,67E_2 - 10,4;$$

$$U_{30} = -0,2E_1 - 4,4.$$

Задаваясь одним из напряжений питания, получаем остальные две величины. Так, задаваясь $E_1 = 20$ В, получаем $E_2 = 18$ В; $U_{30} = -8,8$ В.

Из (14-40) можно заметить, что всегда должно выполняться условие $E_2 > U_0 + \Delta_0$ и что равенство $E_1 = E_2$ (желательное в

¹ В выражениях (14-40), (14-41) символом Δ_0 обозначена «степень насыщения» транзистора T_0 :

$$\Delta_0 = (U_{cn})_0 - [(U_{3n})_0 - U_0].$$

Такой «запас» по насыщению необходим для того, чтобы при положительном синфазном сигнале, когда стоковое напряжение транзистора T_0 уменьшается (по модулю), не было нарушено условие насыщения $U_{cn} > U_{3n} - U_0$.

реальных схемах) возможно только при достаточно больших значениях B_0 , удовлетворяющих условию

$$\sqrt{B_0} > \frac{1}{\sqrt{B_{\text{ср}}} - 1}.$$

В заключение остановимся на напряжении разбаланса U_p . Оно в конечном счете определяется экспериментально, так как теоретически рассчитать индивидуальную асимметрию плеч невозможно. Однако формально оценить влияние тех или иных допусков на величину U_p можно из тех же выражений, которыми описывается статический режим. Снова опуская выкладки, запишем согласно [146] результат:

$$U_p \approx -(E_1 - U_0) \frac{\delta K}{|K_{\text{ср}}|} + \Delta U_{012} - \frac{\Delta U_{034}}{|K_{\text{ср}}|}. \quad (14-42)$$

Здесь $\Delta U_{012} = U_{01} - U_{02}$ и $\Delta U_{034} = U_{03} - U_{04}$ — разбросы пороговых напряжений. Среднее пороговое напряжение U_0 по-прежнему принято одинаковым для всех транзисторов¹.

Из выражения (14-42) следует, что разброс пороговых напряжений активных транзисторов T_1 и T_2 непосредственно определяет величину U_p , а разброс пороговых напряжений нагрузочных транзисторов ослабляется в $K_{\text{ср}}$ раз. Весьма важную роль может играть первое слагаемое, не связанное с разбросом пороговых напряжений. Например, если $E_1 - U_0 = 15$ В, $\delta K = 0,02$ и $K_{\text{ср}} = 5$, то первое слагаемое составляет 60 мВ — значение, вполне сравнимое с разбросом ΔU_0 .

¹ Учет различия пороговых напряжений не принципиален и приводит лишь к более громоздким выражениям для всех рассмотренных выше параметров (см. [146]). Однако при оценке напряжения разбаланса учет разброса величины U_0 , разумеется, принципиально необходим.

Глава пятнадцатая

ТРАНЗИСТОРНЫЕ КЛЮЧИ

15-1. ВВЕДЕНИЕ

Будем называть ключом такую схему, основное назначение которой состоит в замыкании и размыкании цепи нагрузки с помощью управляющих входных сигналов. По аналогии с механическим ключом (контактом) качество транзисторного ключа определяется в первую очередь минимальным падением напряжения на нем в замкнутом состоянии, минимальным током в разомкнутом состоянии, а также скоростью перехода из одного состояния в другое. Ключевые схемы лежат в основе более сложных импульсных схем, поэтому материал данной главы будет широко использован в дальнейшем. Кроме того, некоторые типы транзисторных ключей имеют самостоятельное значение в качестве бесконтактных прерывателей.

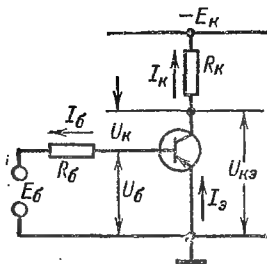


Рис. 15-1. Простейший транзисторный ключ на транзисторе, включенном по схеме ОЭ.

Транзисторный ключ имеет ряд схемных вариантов, простейший из которых приведен на рис. 15-1. Активная нагрузка R_k

включена в коллекторную цепь, а управляющие импульсы поступают от генератора $E_б$ через сопротивление $R_б$, которое, в частности, может быть внутренним сопротивлением генератора. Включение транзистора по схеме ОЭ имеет в импульсной технике столь же широкое распространение, как и в области усилителей. Поэтому такое включение положено в основу последующего анализа, а особенности других включений и других схемных вариантов ключей будут охарактеризованы позднее.

В отличие от усилителей, в которых (за исключением мощных каскадов класса В) транзисторы работают в активном режиме, в ключах (и импульсных устройствах вообще) транзисторы могут работать в нескольких качественно различных режимах. Эти режимы

характеристики полярности напряжений указаны применительно к транзистору *p-n-p*. А именно различают¹:

- 1) режим отсечки ($U_{\text{э}} < 0$; $U_{\text{к}} < 0$);
- 2) нормальный активный режим ($U_{\text{э}} > 0$; $U_{\text{к}} < 0$);
- 3) инверсный активный режим ($U_{\text{э}} < 0$; $U_{\text{к}} > 0$);
- 4) режим насыщения ($U_{\text{э}} > 0$; $U_{\text{к}} > 0$).

Последний было бы правильнее называть режимом двойной инжекции, так как насыщение (тока), как увидим ниже, есть лишь результат такого режима в ключевых схемах.

15-2. СТАТИЧЕСКИЕ ХАРАКТЕРИСТИКИ КЛЮЧА ОЭ

Рассмотрим семейство характеристик транзистора, включенного по схеме ОЭ (рис. 15-2). Проведем на характеристиках линию нагрузки $R_{\text{к}}$. Минимальный ток ключа соответствует точке *A*, а минимальное падение напряжения на ключе — точке *B*. В точке *A* транзистор заперт, так как на его базе действует положительное смещение; в точке *B* транзистор открыт и насыщен.

Рассмотрим подробнее эти два состояния, положив в основу формулы (4-4), описывающие идеализированный транзистор в режиме большого сигнала.

Режим отсечки. Пусть $E_{\text{б}} > 0$ и пусть обе э. д. с. $E_{\text{б}}$ и $E_{\text{к}}$ значительно превышают величину температурного потенциала φ_{T} . Положим, кроме того, что сопротивления $R_{\text{б}}$ и $R_{\text{к}}$ не очень велики, так что падением напряжения на них от «остаточных» токов базы и коллектора можно пренебречь. Тогда можно считать $|U_{\text{э}}| \approx E_{\text{б}} \gg \varphi_{\text{T}}$ и $|U_{\text{к}}| \approx E_{\text{к}} \gg \varphi_{\text{T}}$, т. е. пренебречь экспонентами в формулах (4-4). Переходя в последних от α к β , используя соотношение (4-5) и полагая $\beta_{\text{N}} \gg \beta_{\text{I}}$, получаем следующие значения токов в запертом транзисторе²:

$$I_{\text{э}} = -\frac{1 + \beta_{\text{N}}}{1 + \beta_{\text{I}} + \beta_{\text{N}}} \frac{\beta_{\text{I}}}{\beta_{\text{N}}} I_{\text{к0}} \approx -\frac{\beta_{\text{I}}}{\beta_{\text{N}}} I_{\text{к0}}; \quad (15-1a)$$

$$I_{\text{к}} = \frac{1 + \beta_{\text{N}}}{1 + \beta_{\text{I}} + \beta_{\text{N}}} I_{\text{к0}} \approx I_{\text{к0}}; \quad (15-1б)$$

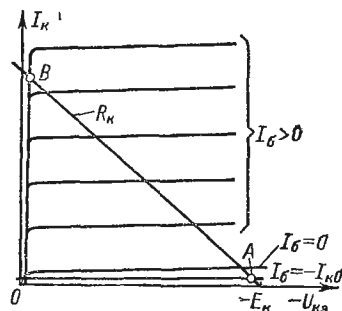


Рис. 15-2. Рабочие точки ключа ОЭ на семействе коллекторных характеристик.

¹ Полярности напряжений указаны применительно к транзистору *p-n-p*.
² Здесь и далее под коэффициентами β понимаются интегральные значения (см. § 4-6), поскольку в ключах и вообще в импульсных схемах транзисторы работают в широком диапазоне токов и напряжений. Напомним, что коэффициенты β зависят от режима; в частности, в режиме отсечки они могут иметь гораздо меньшие значения, чем в нормальном активном режиме.

$$I_6 = -\frac{1+\beta_N}{1+\beta_I+\beta_N} \left(1 + \frac{\beta_I}{\beta_N}\right) I_{K0} \approx -I_{K0}. \quad (15-1в)$$

Отсюда следует, что ток эмиттера значительно меньше остальных двух токов, поскольку $\beta_I \ll \beta_N$; поэтому в режиме отсечки имеем:

$$I_K \approx -I_6 \approx I_{K0}.$$

Рассмотренный случай называют режимом г л у б о к о й отсечки, так как выходной ток имеет минимальное значение. С уменьшением положительного смещения на базе до нуля транзистор продолжает оставаться запертым, но его токи несколько изменяются. Полагая $U_9 \approx E_6 = 0$ в формулах (4-4) и считая по-прежнему

$|U_K| \approx E_K \gg \varphi_T$, получаем приближенные значения токов на границе режима отсечки и активного режима:

$$I_9 \approx \beta_I I_{K0}; \quad (15-2а)$$

$$I_K \approx (1 + \beta_I) I_{K0}; \quad (15-2б)$$

$$I_6 \approx -I_{K0}. \quad (15-2в)$$

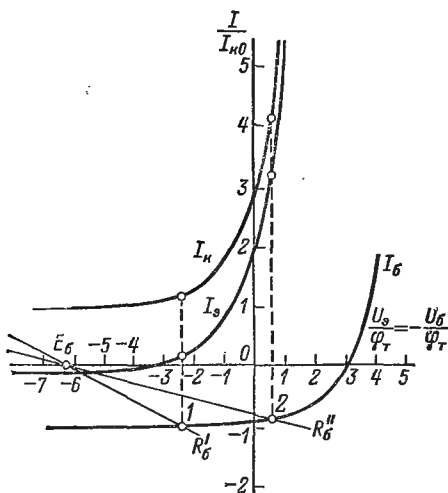


Рис. 15-3. Статические характеристики ключа в режиме отсечки и в начале активного режима.

Сравняя выражения (15-1) и (15-2), приходим к выводу, что ток эмиттера на границе отсечки существенно увеличивается и меняет знак, тогда как ток базы остается практически неизменным. Соответственно ток коллектора возрастает примерно на ту же величину, что и ток эмиттера. Изменение всех трех токов в области отсечки иллюстрируются кривыми на рис. 15-3. Напряжение U_9 , при котором ток I_9 проходит через нуль, нетрудно получить из формулы (4-4а):

$$U_{90} \approx -\varphi_T \ln(1 + \beta_N). \quad (15-3)$$

Если, например, $\beta_N = 20$, то $U_{90} \approx 0,08$ В. Формулы (15-1) и (15-2) были выведены в предположении, что

$$U_{R_6} = I_{K0} R_6 \ll E_6; \quad (15-4а)$$

$$U_{R_K} = I_{K0} R_K \ll E_K. \quad (15-4б)$$

Если сопротивление R_6 велико, так что условие (15-4а) не соблюдается, то токи коллектора и эмиттера могут получиться больше ожидаемых величин.

На рис. 15-3 из точки E_6 проведены две линии нагрузки, соответствующие двум разным сопротивлениям: R_6' и R_6'' ; точки пересечения 1 и 2 определяют действительный режим транзистора. Как видим, при сопротивлении R_6' транзистор еще находится в режиме отсечки, хотя и не такой глубокой, как можно было ожидать, судя по значению E_6 . При сопротивлении же R_6'' транзистор переходит в активный режим, несмотря на положительную величину E_6 . Во избежание такого нежелательного явления следует рекомендовать выбор сопротивления R_6 , исходя из неравенства (15-4а). По аналогичным соображениям сопротивление R_k должно удовлетворять неравенству (15-4б), что обычно легко выполняется. В обоих случаях под током I_{k0} нужно понимать максимальное значение теплового тока, соответствующее наивысшей рабочей температуре.

В заключение напомним, что собственно тепловой ток I_{k0} — это не единственный, а иногда и не главный компонент обратного тока в запертом $p-n$ переходе. Например, у кремниевых транзисторов токи в запертом состоянии определяются токами термогенерации в переходах (см. § 2-6), которые значительно меньше тепловых токов германиевых транзисторов. Кроме того, в обоих типах транзисторов определенную роль могут играть токи утечки, особенно при комнатной и пониженных температурах. Однако в целом режим отсечки в кремниевых транзисторах характерен столь малыми токами, что его теоретический анализ практически не оправдан. Материал данного раздела можно считать относящимся только к германиевым транзисторам.

Спецификой дрейфовых транзисторов в области отсечки является то, что при сравнительно небольших запирающих напряжениях E_6 (1—2, иногда до 5 В) происходит обратимый пробой эмиттерного перехода (см. § 4-12) и ток I_6 резко возрастает; это, однако, не приводит к отпиранию коллекторной цепи.

Режим насыщения. В режиме насыщения оба $p-n$ перехода транзистора смещены в положительном направлении и напряжения на них не превышают 0,35—0,75 В. Обычно э. д. с. E_6 и E_k значительно больше межэлектродных напряжений, поэтому токи I_6 и I_k (а следовательно, и ток I_s) можно считать заданными внешней схемой:

$$I_6 = \frac{E_6 - U_{6э}}{R_6} \approx \frac{E_6}{R_6}; \quad (15-5a)$$

$$I_k = \frac{E_k - U_{кэ}}{R_k} \approx \frac{E_k}{R_k}. \quad (15-5б)$$

В условиях заданных токов критерий насыщения $U_k > 0$; $U_s > 0$ оказывается неудобным, так как требует предварительного определения напряжений. Поэтому целесообразно (во всяком случае в статическом режиме) иметь критерий насыщения, выраженный через токи. Такой критерий легко получить из выражения (4-6). А именно, когда транзистор находится в активном режиме ($U_k < 0$), имеем $I_k > \alpha I_s$. На границе насыщения ($U_k = 0$) имеем $I_k = \alpha I_s$.

Наконец, в режиме насыщения, когда $U_{\kappa} > 0$, имеем $I_{\kappa} < \alpha I_{\text{э}}$, откуда получаем критерий насыщения в виде:

$$I_{\text{э}} > \frac{I_{\kappa}}{\alpha} = I_{\text{э.н}}, \quad (15-6)$$

где $I_{\text{э.н}}$ — насыщающий ток эмиттера.

Поскольку коэффициент α характерен для включения ОБ, а мы рассматриваем ключ ОЭ, нужно в неравенстве (15-6) перейти от коэффициента α к коэффициенту β ; тогда критерием насыщения будет условие

$$I_{\text{б}} > \frac{I_{\kappa}}{\beta} = I_{\text{б.н}}, \quad (15-7)$$

где $I_{\text{б.н}}$ — насыщающий ток базы.

Как видим, режим насыщения определяется не значением токов, а их соотношением и может иметь место при весьма малых токах (порядка микроампер и меньше).

Опыт показывает, что удобно иметь не только критерий насыщения, но и количественную характеристику «глубины» насыщения. Такой параметр называется *степенью насыщения* [147] и определяется как относительное превышение базовым током $I_{\text{б}}$ того значения $I_{\text{б.н}}$, которое характерно для границы насыщения¹:

$$N = \frac{I_{\text{б}} - I_{\text{б.н}}}{I_{\text{б.н}}} = \frac{\beta I_{\text{б}} - I_{\kappa.н}}{I_{\kappa.н}}. \quad (15-8)$$

Рассмотрим теперь зависимость межэлектродных напряжений насыщенного транзистора от заданных токов. Используя выражения (4-4), находим:

$$U_{\text{эб}} = \varphi_T \ln \left[\frac{(1 + \beta_I) I_{\text{б}} + I_{\kappa}}{I_{\text{э0}}^*} + 1 \right]; \quad (15-9a)$$

$$U_{\kappa\text{б}} = \varphi_T \ln \left[\frac{\beta_N I_{\text{б}} - I_{\kappa}}{I_{\kappa0}^*} + 1 \right], \quad (15-9б)$$

где $I_{\text{э0}}^* = (1 + \beta_I) I_{\text{э0}}$ — величина, аналогичная току $I_{\kappa0}^*$ [см. (4-70)]. Напряжение $U_{\text{бэ}} = -U_{\text{эб}}$ есть входное напряжение ключа. Что касается «остаточного» напряжения $U_{\kappa\text{э}}$ в цепи нагрузки, то оно определяется разностью $U_{\kappa\text{б}} - U_{\text{эб}}$ и с учетом (4-5) может быть приведено к виду

$$U_{\kappa\text{э}} = \varphi_T \ln \left(\alpha_I \frac{1 - I_{\kappa}/\beta_N I_{\text{б}} + I_{\kappa0}/\alpha_N I_{\text{б}}}{1 + I_{\kappa}/(1 + \beta_I) I_{\text{б}} + I_{\text{э0}}/I_{\text{б}}} \right). \quad (15-10a)$$

¹ В работе [147], автор которой ввел понятие степени насыщения, последняя обозначается буквой S и определяется как $I_{\text{б}}/I_{\text{б.н}}$, что, на наш взгляд, менее удобно.

Если $I_6 \gg I_{к0}, I_{э0}$, а $\beta_N I_6 > (3 \div 4) I_k$ (т. е. степень насыщения $N > 2 \div 3$), то «остаточное» напряжение выражается более простой формулой¹:

$$U_{кэ} \approx \varphi_T \ln \frac{\alpha_I}{1 + I_k / (1 + \beta_I) I_6}. \quad (15-10б)$$

Из выражений (15-9), (15-10) и рис. 15-4 можно сделать следующие выводы.

1. С увеличением базового тока напряжение на эмиттерном переходе $U_{эб}$ меняется мало, напряжение на коллекторном переходе $U_{кб}$ растет, а напряжение $U_{кэ}$ (т. е. напряжение на «замкнутом» ключе) уменьшается. В пределе, при $I_6 \rightarrow \infty$, получаем:

$$U_{кэ, \text{ мин}} = \varphi_T \ln \alpha_I = -\varphi_T \ln \left(1 + \frac{1}{\beta_I}\right) \approx -\frac{\varphi_T}{\beta_I}. \quad (15-10в)$$

Например, при $\beta_I = 0,5 \div 5$ напряжение $U_{кэ, \text{ мин}}$ составляет соответственно -50 и -5 мВ.

2. Напряжения на переходах $U_{эб}$ и $U_{кб}$ зависят от температуры примерно так же, как и в диодах (см. § 2-8). Однако температурные чувствительности каждого из этих напряжений несколько различны (из-за различия в токах и коэффициентах β) и зависят от степени насыщения. Что касается напряжения $U_{кэ}$, то оно, будучи разностью напряжений на переходах, зависит от температуры гораздо слабее. Это подтверждается формулами (15-10).

Так, если пренебречь влиянием тепловых токов и положить $I_6 \geq 0,1 I_k$, то из (15-10б) при $\beta_I > 2$ нетрудно получить, что температурная чувствительность напряжения $U_{кэ}$ составляет не более $-0,15$ мВ/°С, т. е. по крайней мере на порядок меньше, чем температурная чувствительность отдельных переходов.

3. Начиная от значений $N = 3 \div 5$ и выше, межэлектродные напряжения мало зависят от тока базы, поэтому более высокие степени насыщения обычно нецелесообразны.

4. У кремниевых транзисторов, у которых тепловые токи на несколько порядков меньше, чем у германиевых, напряжения на переходах получаются больше. Напряжение $U_{кэ}$ мало зависит от тепловых токов и потому примерно одинаково у обоих типов тран-

¹ Следует заметить, что интегральные коэффициенты β_N и β_I в режиме насыщения не соответствуют формуле (4-73), которая записана для активного режима, т. е. для условия $U_k < 0$. Исходя из выражения (4-6) и аналогичного выражения для инверсного включения, можно показать, что в режиме насыщения

$$\beta_N = \frac{I_k + I_{к0} (e^{U_k / \varphi_T} - 1)}{I_6 - I_{к0} (e^{U_k / \varphi_T} - 1)}; \quad \beta_I = -\frac{I_э - I_{э0} (e^{U_э / \varphi_T} - 1)}{I_6 - I_{э0} (e^{U_э / \varphi_T} - 1)},$$

где для инверсного включения следует считать $I_э < 0$.

зисторов (если не учитывать омических падений напряжения — см. ниже).

Сравнительно малые межэлектродные напряжения насыщенного транзистора иногда позволяют считать его «эквипотенциальной

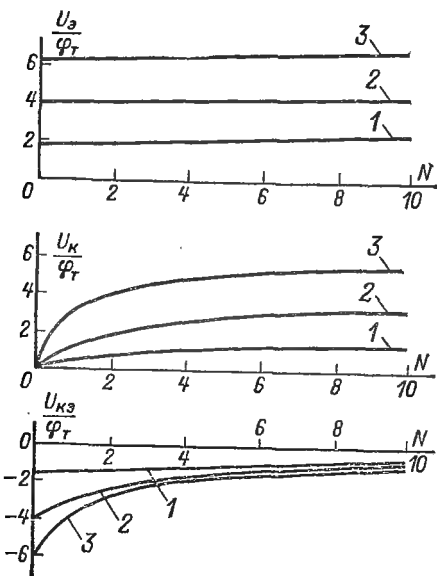


Рис. 15-4. Зависимость межэлектродных напряжений насыщенного ключа от токов и степени насыщения.

1 — $I_{\text{к}} = 10 I_{\text{к}0}$; 2 — $I_{\text{к}} = 10^2 I_{\text{к}0}$; 3 — $I_{\text{к}} = 10^3 I_{\text{к}0}$.

точкой», что существенно упрощает статический анализ многих импульсных схем. На рис. 15-5, а в качестве примера показана обобщенная схема транзисторного ключа с источниками смещения и сопротивлениями в цепях всех трех электродов. В режиме насыщения и при условии $E_{\text{э}}, E_{\text{б}}, E_{\text{к}} > 1$ В схеме на рис. 15-5, а можно заменить эквивалентной схемой на рис. 15-5, б, в которой потенциал узловой точки — потенциал транзистора — определяется известной формулой узлового напряжения [62]:

$$U_T = \frac{\sum E_i g_i}{\sum g_i}, \quad (15-11a)$$

где g_i — проводимости ветвей, а E_i — э. д. с. в этих ветвях.

В случае кремниевых транзисторов целесообразно

вместо $E_{\text{б}}$ подставлять $E_{\text{б}} + 0,7$ В, чтобы учесть падения напряжения на эмиттерном и коллекторном переходах. В этом случае $U_T \approx U_{\text{э}} \approx U_{\text{к}}$, а потенциал базы $U_{\text{б}} \approx U_T - 0,7$ В.

В более общих случаях (рис. 15-6), когда в схеме дополнительно действует генератор тока, формула (15-11а) также обобщается и для рис. 15-6, а и б принимает соответственно вид:

$$U_T = \frac{\sum E_i g_i + I}{\sum g_i}; \quad (15-11б)$$

$$U_T = \frac{\sum E_i g_i + I \frac{g_m}{g_{m1}}}{\sum g_i}, \quad (15-11в)$$

где g_{m1} — та часть проводимости g_m , которая шунтирует генератор тока.

Формула (15-11в) остается в силе и в том случае, когда генератор тока подключен не к «земле», как на рис. 15-6, б, а к узловой точке. В этом случае под проводимостью g_{m1} следует опять понимать ту часть проводимости g_m , которая шунтирует генератор (на рис. 15-6, б — это проводимость g_{m2}).

Выше мы не учитывали объемных сопротивлений слоев. Учет этих сопротивлений сводится к тому, что межэлектродные напряжения следует изменить на величину падения напряжения в соответствующем слое, т. е. (см. рис. 4-12, а)

$$|U_{бэ}| = U_э + I_э r_{ээ} + I_б r_б; \quad (15-12а)$$

$$|U_{кэ}| = |U_к - U_э| + I_э r_{ээ} + I_к r_{кк}, \quad (15-12б)$$

где $r_{ээ}$ и $r_{кк}$ — сопротивления слоев; $U_э$ и $U_к$ — напряжения на переходах, вычисляемые по формулам (15-9).

Сопротивление $r_б$ оказывается существенным практически всегда. Напротив, сопротивление $r_{ээ}$ приходится учитывать лишь в исключительных случаях, поскольку слой эмиттера наиболее низкоомный. Что касается сопротивления $r_{кк}$, то его роль в значительной степени зависит от типа транзистора. У сплавных транзисторов оно обычно столь же мало, как и $r_{ээ}$ (не более одного ома).

У планарно-эпитаксиальных транзисторов оно значительно больше и может составлять десятки и сотни ом. При этом суммарное напряжение $|U_{кэ}|$ на «замкнутом» ключе может при сильных токах достигать 1—1,5 В. Такое большое значение не идет ни в какое сравнение с величиной, определяемой формулой (15-10в) для идеализированного транзистора. При расчетах по формулам (15-11) сопротивление $r_{кк}$ следует складывать с $R_к$. Найденное значение U_T близко к потенциалу $U_э$, а потенциал $U_к$ получается путем добавления члена $-I_к r_{кк}$. В случае кремниевых транзисторов в расчетах потенциалов $U_T \approx U_э \approx U_к$ и $U_б$ учитываются также напряжения на открытых переходах (см. с. 466 перед петитом).

Рис. 15-6. Обобщенные схемы, соответствующие теореме узловых напряжений.

а — генератор тока подключен к узлу; б — генератор тока подключен к части проводимости.

Дифференцируя (15-12а) и (15-12б) соответственно по $I_б$ и $I_к$, получаем дифференциальные входное и выходное сопротивления ключа в режиме насыщения:

$$R_{вх.н} = r_{бэ} + r_{ээ} + r_б; \quad (15-13а)$$

$$R_{вых.н} = r_{кэ} + r_{ээ} + r_{кк}. \quad (15-13б)$$

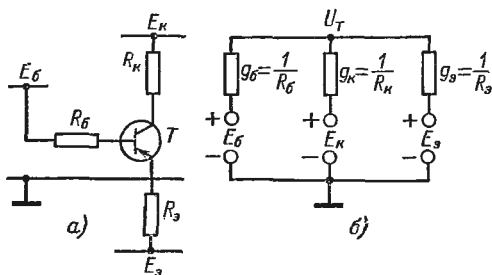
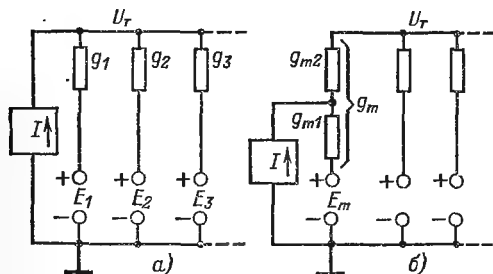


Рис. 15-5. Общий случай транзисторного ключа (а) и его эквивалентная схема в режиме насыщения (б).



В этих выражениях первые слагаемые в правых частях являются дифференциальными сопротивлениями переходов в режиме насыщения и рассчитываются как производные от (15-9):

$$r_{6э} \approx \frac{\varphi_T}{(1 + \beta_I) I_6 + I_K}; \quad (15-14a)$$

$$r_{кэ} \approx \frac{\varphi_T I_6 (1 + \beta_I + \beta_N)}{(\beta_N I_6 - I_K) [(1 + \beta_I) I_6 + I_K]}. \quad (15-14б)$$

Оба эти сопротивления монотонно уменьшаются с ростом тока I_6 . С ростом тока I_K сопротивление $r_{6э}$ тоже монотонно уменьшается, а сопротивление $r_{кэ}$ имеет минимум при токе $I_{K, \text{опт}} = 0,5 (\beta_N - \beta_I - 1) I_6$:

$$r_{кэ, \text{ мин}} = \frac{4\varphi_T}{I_6 (1 + \beta_I + \beta_N)}. \quad (15-14в)$$

В области малых коллекторных токов ($I_K \ll \beta_I I_6$) сопротивление $r_{кэ}$ имеет величину:

$$r_{кэ 0} \approx \varphi_T / (1 + \beta_I) I_6. \quad (15-14г)$$

Расчеты показывают, что дифференциальные сопротивления переходов в режиме насыщения обычно лежат в пределах нескольких ом. Такое же и даже меньшее значение имеет омическое сопротивление $r_{ээ}$. Поэтому сопротивлением насыщенного ключа $R_{\text{вых.н}}$ можно либо пренебрегать, либо (в случае планарно-эпитаксиальных, сплавно-диффузионных и мезатранзисторов) считать его равным $r_{кк}$. Что касается входного сопротивления $R_{\text{вх.н}}$, то оно близко к сопротивлению r_6 и этим часто пользуются при анализе насыщенных ключей, полагая

$$R_{\text{вх. н}} \approx r_6. \quad (15-15)$$

Следует, однако, иметь в виду, что значение r_6 в режиме насыщения может существенно (в несколько раз) отличаться от значения, соответствующего усиленному режиму транзистора. Причина различия состоит в модуляции базового сопротивления накапливающимися носителями (рис. 15-7). Точно рассчитать уменьшение r_6 под действием модуляции трудно, так как распределение заряда в базе насыщенного транзистора имеет сложный характер. Приближенный анализ (см. [148], а также 2-е издание данной книги, стр. 384) приводит к соотношению вида

$$r_6 \approx r_{61} + \frac{r_{62}}{1 + I_6 / I^*}, \quad (15-16)$$

которое качественно подтверждается измерениями [149]. Ток I^* у маломощных транзисторов может составлять доли миллиампера, так что модуляция сопротивления r_6 нередко становится существенной уже при токах базы порядка 1 мА.

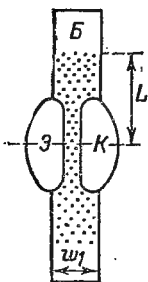


Рис. 15-7. Симметричный транзистор со ступенчатым распределением заряда, модулирующего сопротивление базы.

Использование прерывателей в усилителях постоянного тока было описано в § 13-6. В таких прерывателях (рис. 15-8) транзисторы работают в ключевом режиме, имеющем, однако, особенности по сравнению с описанным выше: во-первых, источником питания служит источник сигнала $U_{вх}$ (т. е. «напряжение питания» не остается постоянным); во-вторых, напряжение $U_{вх}$, играющее роль E_b , может иметь весьма малое значение (десятки милливольт и меньше).

Пусть ключ на рис. 15-8, а заперт по базе напряжением $+E_b$. Тогда при всех значениях $U_{кз} < E_b$ в выходной цепи протекает ток, близкий к $I_{к0}$, и соответствующая характеристика идет почти

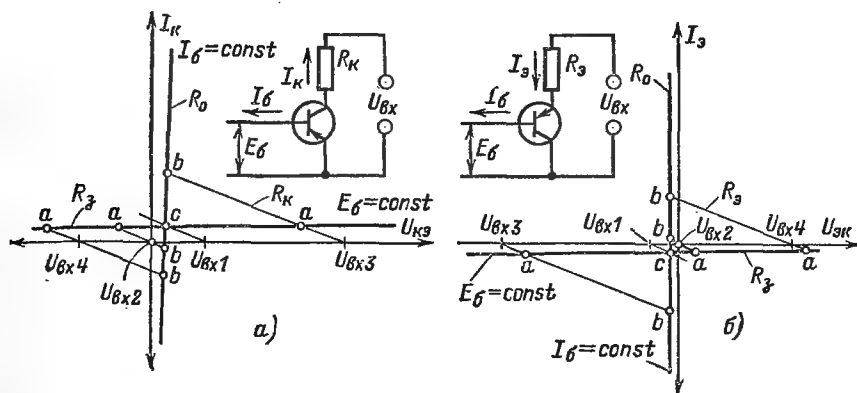


Рис. 15-8. Рабочие характеристики прерывателя.

а — по схеме ОЭ; б — по схеме «ОК» ($U_{вх} = U_{аб}$).

горизонтально. Пусть теперь в цепи базы задан положительный ток I_6 . Тогда при всех токах $I_K < \beta_N I_6$ транзистор насыщен, выходное напряжение очень мало и соответствующая характеристика идет почти вертикально. Назовем характеристику, по которой перемещаются точки а (рис. 15-8, а), *линией запираения*, а характеристику, по которой перемещаются точки б, — *линией отпираания*. Эти линии, показанные на рис. 15-8, а в области, близкой к началу координат, вообще говоря, свойственны любому ключу (на рис. 15-2 на этих линиях расположены точки А и В).

Из рис. 15-8, а видно, что в отличие от идеального (контактного) прерывателя, у которого линии запираения и отпираения совпадали бы с осями координат, в транзисторном ключе эти линии имеют конечный наклон, а главное, их точка пересечения не совпадает с началом координат. Последнее обстоятельство приводит к тому, что при конечном сигнале $U_{вх1}$ получается нулевое выходное напряжение и, наоборот, при нулевом сигнале $U_{вх2} = 0$ получается ко-

нечное выходное напряжение¹. Если бы точка c совпадала с началом координат, такого несоответствия не было бы, но величины $U_{\text{вых}}$ и $U_{\text{вх}}$ все же различались бы из-за наклона характеристик. Таким образом, транзисторному прерывателю свойственны два вида погрешностей: *погрешность сдвига* и *погрешность наклона*. Очевидно, что оба типа погрешностей имеют тем меньшее значение, чем больше сигнал (см. случаи $U_{\text{вх}3}$ и $U_{\text{вх}4}$ на рис. 15-8, а). Если выполняется условие $|U_{\text{вх}}| \gg U_{\text{вх}1}$, то погрешность прерывателя невелика, причем положительным $U_{\text{вх}}$ соответствует отрицательная погрешность ($U_{\text{вых}} < U_{\text{вх}}$), а отрицательным $U_{\text{вх}}$ — положительная погрешность ($U_{\text{вых}} > U_{\text{вх}}$).

Для количественной оценки погрешностей нужно знать координаты точки c и дифференциальные сопротивления обеих характеристик. Считая, что наклон линии запириания соответствует некоторому сопротивлению R_z (которое учитывает утечки, ток термогенерации в переходе и другие факторы), получаем:

$$I_c = I_{k0} + E_6/R_z. \quad (15-17)$$

Напряжение в точке c можно найти, подставив $I_k = I_c$ в формулу (15-10а):

$$U_c = \varphi_T \ln \frac{\alpha_I (I_6 + I_{k0}/\alpha_N - I_c/\beta_N)}{I_6 + I_{s0} + I_c/(1 + \beta_I)}. \quad (15-18a)$$

В практических случаях ток базы достаточно велик, так что можно пренебречь вторым и третьим слагаемыми в числителе и знаменателе (15-18а). Тогда получается весьма простое выражение, совпадающее с (15-10в):

$$|U_c| \approx \frac{\varphi_T}{\beta_I}. \quad (15-18б)$$

Наклон линии запириания характеризуется принятым выше сопротивлением R_z (обычно не меньше 1 МОм), а наклон линии отпириания — сопротивлением $R_o = R_{\text{вых.н}}$ [см. (15-13б)]. Обычно R_o составляет несколько ом.

Необходимо иметь в виду, что параметры I_c , U_c , R_z и R_o существенно зависят от температуры. Расчет этих зависимостей в принципе несложен, если известны функции $\beta(T)$ и $R_z(T)$. С увеличением температуры следует ожидать увеличения тока I_c (особенно у германиевых транзисторов), некоторого уменьшения напряжения U_c , а также уменьшения обоих сопротивлений. Все эти изменения вызывают изменения модулированного напряжения при неизменном входном сигнале и, следовательно, равносильны дрейфу последнего.

Большое распространение в прерывателях имеет инверсное включение транзистора (рис. 15-8, б), которое по сравнению с нор-

¹ На рис. 15-8, а и б выходное напряжение есть разность абсцисс точек a и b , в которых линия нагрузки пересекается с характеристиками $E_6 = \text{const}$ и $I_6 = \text{const}$.

малым включением обеспечивать меньшие ток I_C и напряжение U_C . Такой ключ часто обозначают буквами ОК, хотя в данном случае эмиттерный переход работает с отрицательным смещением, что не соответствует определению включения ОК (рис. 4-3). Поэтому ниже мы будем ставить обозначение ОК в кавычки.

В ключах «ОК» ток I_C есть ток эмиттера при запертых переходах, когда на базу задано положительное смещение E_6 , а сигнал $U_{эк} < E_6$. Используя выражение (15-1а) и вводя, как в схеме ОЭ, сопротивление R_3 , получаем:

$$I_C = -\frac{\beta_I}{\beta_N} I_{к0} - \frac{E_6}{R_3}. \quad (15-19)$$

Напряжение U_C можно найти, заменив в формуле (15-10а) ток I_K на $I_э - I_6$ и подставив $I_э = I_C$, где $I_C < 0$. Кроме того, поскольку $U_{эк} = -U_{кэ}$, нужно изменить знак в формуле (15-10а). Тогда

$$U_C = \varphi_T \ln \frac{\alpha_N (I_6 + I_{э0} / \alpha_I + I_C / \beta_I)}{I_6 + I_{к0} - I_C / (1 + \beta_N)}. \quad (15-20а)$$

Практически, как и в схеме ОЭ, можно пренебречь последними двумя членами в числителе и знаменателе и получить упрощенное выражение:

$$|U_C| \approx \frac{\varphi_T}{\beta_N}. \quad (15-20б)$$

Поскольку $\beta_N \gg \beta_I$, напряжение U_C в ключе «ОК» получается по крайней мере на порядок меньше, чем в ключе ОЭ, что, конечно, является большим преимуществом первого. Следует, однако, иметь в виду, что значения β в области малых токов могут существенно отличаться от номинальных, особенно у кремниевых транзисторов (§ 4-6). Сопротивления R_3 и R_0 в ключе «ОК» можно считать такими же, как в ключе ОЭ.

Формулы (15-18) и (15-20) не учитывают падения напряжения в сопротивлениях слоев $r_{ээ}$ (для нормального включения) или $r_{кк}$ (для инверсного включения «ОК»). Между тем эти падения напряжения (особенно на сопротивлении $r_{кк}$) могут быть существенными, поскольку через указанные сопротивления протекает не только сравнительно малый выходной ток, но и значительно больший управляющий ток базы. Таким образом, более точными выражениями будут:

$$|U_C| = \frac{\varphi_T}{\beta_I} + I_6 r_{ээ} \quad (15-21а)$$

Для нормального включения и

$$|U_C| = \frac{\varphi_T}{\beta_N} + I_6 r_{кк} \quad (15-21б)$$

Для инверсного включения («ОК»).

Пусть, например, используется транзистор с параметрами $\beta_N = 50$, $\beta_I = 5$, $r_{ээ} = 1$ Ом, $r_{кк} = 10$ Ом и пусть $I_6 = 1$ мА. Тогда $\varphi_T/\beta_I = 5$ мВ, $I_6 r_{ээ} = 1$ мВ, $\varphi_T/\beta_N = 0,5$ мВ и $I_6 r_{кк} = 10$ мВ. Как видим, при нормальном включении «омическая поправка» не очень существенна, тогда как при инверсном включении она является основной составляющей остаточного напряжения. В результате последнее оказывается для инверсного включения («ОК») больше, чем для нормального (ОЭ).

Из выражений (15-21) ясно видно, что напряжение U_C растет с увеличением тока I_6 . Однако в области малых базовых токов напряжение U_C растет с уменьшением тока I_6 , так как при этом транзистор постепенно выходит из насыщения [см. формулы (15-18а) и (15-20а)]. Таким образом, функция $U_C(I_6)$ имеет минимум. В принципе этот минимум нетрудно найти, добавляя в правые части (15-18а) и (15-20а) соответственно члены $-I_6 r_{ээ}$ и $-I_6 r_{кк}$ и дифференцируя по I_6 . В общем случае выражения получаются громоздкими. Они упрощаются, если пренебречь тепловыми токами (поскольку в прерывателях чаще используются кремниевые транзисторы) и положить $\beta_N \gg \beta_I + 1$ и $I_6 \gg I_C$. Тогда для нормального включения оптимальный ток базы имеет вид:

$$I_{6, \text{ опт}} \approx \sqrt{\frac{\varphi_T I_C}{(\beta_I + 1) r_{ээ}}}. \quad (15-22)$$

Для инверсного включения нужно заменить $r_{ээ}$ на $r_{кк}$ и поменять индекс при коэффициенте β .

Очевидно, что оптимальный ток при инверсном включении будет меньше, чем при прямом. Обычно значение $I_{6, \text{ опт}}$ лежит в пределах от долей миллиампер до нескольких миллиампер. Заметим, что формула (15-22) является весьма приближенной, так как при ее выводе не учитывались зависимость $\beta(I_6)$, а также особенности высокого уровня инжекции. Последний фактор проанализирован в статье [150]. Результатом является замена линейного соотношения $U_C \sim I_6$ в области больших токов [см. (15-21)] соотношением $U_C \sim \sqrt{I_6}$.

Для прерывателей, используемых в усилителях постоянного тока, исключительную роль играет стабильность точки c , в первую очередь стабильность напряжения U_C (ток I_C для кремниевых транзисторов несуществен и его нестабильностью можно пренебречь). В § 15-2 отмечалось, что температурный коэффициент напряжения $U_{кв}$ очень мал, если ток I_6 заметно превышает тепловые токи. При этом, однако, не учитывались зависимости $\beta(T)$ и $r_{кк}(T)$ (последняя — для инверсного включения «ОК»). Учитывая, что параметры β и $r_{кк}$ увеличиваются с ростом температуры, приходим к выводу, что результирующая температурная чувствительность может иметь разные знаки в зависимости от тока базы, поскольку от него зависят значения β и $I_6 r_{кк}$. Этот вывод подтверждается на практике [151]: при токах базы, близких к оптимальному значению (15-22), температурная чувствительность может составлять всего несколько милливольт на градус, тогда как при больших и особенно меньших

тока она доходит до десятков и сотен микровольт на градус. Таким образом, в диапазоне $\pm 60^\circ \text{C}$ можно добиться (подбором транзистора и режима) температурного дрейфа около 100 мкВ.

Не меньшую роль для усилителей постоянного тока играет временная стабильность напряжения U_C . В этом отношении приходится ориентироваться только на эксперимент, который показывает [151], что временной дрейф имеет индивидуальный характер у разных типов и экземпляров транзисторов. Как правило, в течение первых 1—2 ч работы дрейф может составить несколько милливольт, а в дальнейшем он лежит в пределах 50—200 мкВ.

Погрешность сдвига можно существенно (в 5—10 раз) уменьшить и, следовательно, повысить чувствительность прерывателя, если использовать балансные схемы [148, 121]. Пример такой схемы показан на рис. 15-9. Принцип компенсации состоит в том, что в отсутствие входного сигнала ($U_{вх} = 0$) потенциалы U_P и U_Q одинаковы и $U_{вых} = 0$, как и требуется для устранения погрешности сдвига. Равенство $U_P = U_Q$ возможно при условиях $I_{э1} = -I_{э2}$ и $U_{э1} = -U_{э2}$.

Поскольку один из транзисторов закрыт, а другой открыт, равенство модулей и одновременно различие в знаках величин $I_{э}$ и $U_{э}$ могут быть обеспечены только тогда, когда рабочая точка первого транзистора лежит на линии запираения в IV квадранте, а рабочая точка второго — на линии отпираения во II квадранте (см. рис. 15-8, б)¹.

Суммарные сопротивления $R_r + R_s$ и $R'_r + R'_s$ должны быть одинаковыми, а их значение выбирается таким, чтобы линия нагрузки $R_r + R_s$, проведенная из начала координат, обеспечила равенство $|U_a| = |U_b|$ или $|I_a| = |I_b|$ на рис. 15-8. Такой выбор осуществляется графически или по приближенной формуле:

$$R_s + R_r \approx U_C / I_C;$$

последняя получена из графического построения при условиях

$$U_C / R_s \ll I_C; \quad U_C / R_o \gg I_C.$$

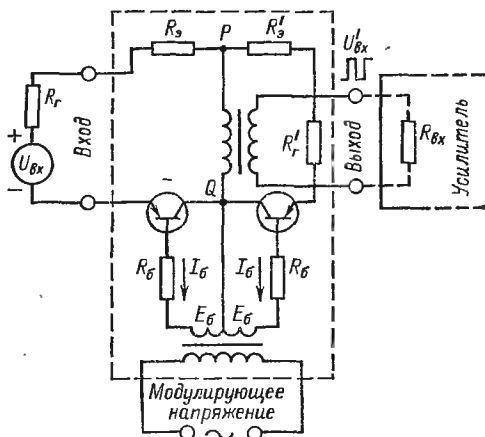


Рис. 15-9. Однотактный компенсированный прерыватель.

¹ Для схемы ОЭ (см. рис. 15-8, а) рабочими квадрантами характеристик будут соответственно II и IV.

Переходные процессы в прерывателях подчиняются общим законам динамики транзисторных ключей (см. § 15-5) и потому отдельно не рассматриваются. Некоторые особенности переходных процессов, а также дополнительные сведения о прерывателях можно найти в монографии [151].

15-4. МЕТОД ЗАРЯДА

Прежде чем анализировать переходные процессы в транзисторном ключе, рассмотрим метод, с помощью которого такой анализ проводится гораздо проще, чем при использовании уравнений непрерывности. В основе этого метода, называемого методом заряда [153, 154], лежит условие нейтральности базы (см. § 1-7). Согласно этому условию в любой точке базы положительный и отрицательный заряды одинаковы и изменяются с одинаковой скоростью. В базе типа n положительный заряд обусловлен донорами и дырками, а отрицательный — только электронами. Следовательно, исходя из условия нейтральности, можно записать:

$$Q_n + Q_p = Q_n.$$

Дифференцируя обе части по времени, получаем:

$$\frac{dQ_n}{dt} + \frac{dQ_p}{dt} = \frac{dQ_n}{dt}. \quad (15-23)$$

Здесь каждое слагаемое имеет размерность тока, и задача состоит в том, чтобы установить связь этих слагаемых с теми токами, которые задаются при расчете или могут быть вычислены.

Изменение заряда доноров Q_n связано с изменением толщины базы или, что то же самое, ширины переходов. Иначе говоря, первое слагаемое равно сумме токов, перезаряжающих барьерные емкости:

$$\frac{dQ_n}{dt} = C_n \frac{dU_n}{dt} + C_k \frac{dU_k}{dt}.$$

Изменение электронного заряда Q_n обусловлено тремя причинами: базовым током, рекомбинацией в базе и электронными составляющими токов I_n и I_k . Последние всегда отрицательны, так как в процессе инжекции электроны уходят из базы. Рекомбинационный ток тоже отрицателен, так как рекомбинация приводит к уменьшению числа электронов. Поскольку речь идет о рекомбинации неравновесных электронов, из формулы (1-51б), интегрируя по объему, легко получить для этого тока значение $-\Delta Q_n/\tau$, где ΔQ_n — неравновесный заряд электронов. Тогда в целом¹

$$\frac{dQ_n}{dt} = I_b - I_n - \frac{\Delta Q_n}{\tau}.$$

¹ Как видно, в стационарном режиме ($dQ_n/dt = 0$) ток базы состоит из двух компонентов, один из которых обусловлен рекомбинацией, а второй — электронными потоками через переходы. В нестационарном режиме добавляется третий компонент (dQ_n/dt), обусловленный накоплением электронов в базе.

Здесь I_n — сумма электронных токов через переходы, причем токи могут быть обусловлены как инжекцией электронов в смежный слой, так и рекомбинацией в области переходов. Подставляя полученные значения производных в (15-23) и учитывая, что $\Delta Q_n = \Delta Q_p$, получаем следующее дифференциальное уравнение для неравновесного дырочного заряда:

$$C_a \frac{dU_a}{dt} + C_k \frac{dU_k}{dt} + \frac{dQ_p}{dt} + \frac{\Delta Q_p}{\tau} = I_6 - I_n. \quad (15-24a)$$

Первый член в левой части, а также ток I_n в правой части не всегда существенны. Поэтому практически пользуются уравнением

$$C_k \frac{dU_k}{dt} + \frac{dQ}{dt} + \frac{Q}{\tau} = I_6 \quad (15-24б)$$

и еще более упрощенным его вариантом

$$\frac{dQ}{dt} + \frac{Q}{\tau} = I_6, \quad (15-24в)$$

в которых опущены индекс p и знак Δ .

Удобство выражений (15-24) заключается в том, что они в отличие от уравнений непрерывности представляют собой обыкновенные дифференциальные уравнения, решать которые проще: нужно знать только начальные условия.

Практическое использование *метода заряда* подразумевает еще наличие простой связи между тем «внутренним» параметром — зарядом, который получается из приведенных уравнений, и таким «внешним» параметром, как коллекторный ток, который обычно представляет главный интерес при анализе.

В активном режиме транзистора связь между неравновесным зарядом Q и током I_k содержится в выражениях (4-62) и (4-136). Запишем их в следующем общем виде:

$$Q = \frac{t_D}{\alpha(1+\eta)} I_k, \quad (15-25a)$$

где $\eta = 0$ для бездрейфовых транзисторов.

Переходя от времени пролета $t_D/1 + \eta$ к времени жизни τ , можно записать (15-25a) в другой форме¹:

$$Q \approx \frac{\tau}{\beta} I_k. \quad (15-25б)$$

Эта форма соответствует наиболее распространенной схеме ОЭ с усилительным параметром — коэффициентом β . Выражения (15-25) основаны на допущении о квазистационарном распределении заряда в базе; во время коротких переходных процессов они дают большую погрешность. Поскольку любые возмущения

¹ Связь между временем диффузии (в общем случае — временем пролета) и временем жизни вытекает из соотношений (4-85) и (4-82).

концентраций, происходящие вблизи эмиттера, распространяются до коллектора за время порядка τ_a , можно считать формулы (15-25) действительными тогда, когда длительность переходных процессов превышает время распространения носителей вдоль базы.

Для решения уравнений (15-24) необходимо знать начальные значения заряда. Обычно эти начальные значения соответствуют стационарным условиям, при которых все производные по времени равны нулю. Тогда получаем:

$$Q = I_{\epsilon} \tau, \quad (15-26)$$

т. е. в стационарном режиме **н е р а в - н о в е с н ы й** заряд дырок пропорционален базовому току. Это важное соотношение действительно не только в активной области, но и в области насыщения.

На границе активной области и области насыщения, когда еще действительны формулы (15-25), **г р а н и ч н ы й** заряд имеет величину:

$$Q_{гр} \approx \frac{\tau}{\beta} I_{к.н}, \quad (15-27)$$

Рис. 15-10. Распределение дырок в базе при активном, граничном и насыщенном режимах.

где $I_{к.н}$ — ток насыщения. Выражение (15-27), как и соотношения (15-25), действительно только при не очень коротких переходных процессах. Величина граничного заряда широко используется при анализе насыщенных ключей в качестве критерия перехода их из активной области в область насыщения и обратно.

Выразив токи в формуле (15-8) через заряды (15-26) и (15-27), получим наглядную интерпретацию степени насыщения:

$$N = \frac{Q - Q_{гр}}{Q_{гр}}.$$

Разность $Q - Q_{гр} = Q_{изб}$ называют **избыточным зарядом**; следовательно, степень насыщения характеризует относительную величину избыточного заряда. На рис. 15-10 показано распределение концентрации дырок (и тем самым заряда) для разных режимов работы транзистора.

В тех случаях, когда переходные процессы анализируются операторным методом, следует записать уравнение заряда в операторной форме и найти изображение заряда. Воспользуемся сначала простейшим уравнением (15-24в). Подставляя общее соотношение $dQ/dt \doteq s [Q - Q(0)]$, получаем:

$$Q(s) = \frac{I_{\epsilon}(s) \tau + (s\tau) Q(0)}{1 + s\tau}. \quad (15-28a)$$

В частном случае, когда $Q(0) = 0$, получаем:

$$Q(s) = \frac{I_6(s)\tau}{1 + s\tau}. \quad (15-286)$$

Отсюда, зная функцию $i_6(t)$ и ее изображение $I_6(s)$, можно найти по таблицам оригинал $Q(t)$ и затем с помощью (15-25) ток $i_k(t)$.

Уравнение (15-24в) позволяет ввести понятие *сильного управляющего сигнала*, которое часто упрощает анализ ключевых схем. Пусть сначала ключ находился в стационарном режиме ($dQ/dt = 0$), а затем ток базы скачком изменился на величину ΔI_6 . Поскольку заряд Q не может измениться мгновенно, получаем в первый момент: $(dQ/dt)_0 = \Delta I_6$; следовательно, заряд сначала меняется линейно:

$$\Delta Q(t) \approx \left(\frac{dQ}{dt}\right)_0 t = \Delta I_6 t, \quad (15-29a)$$

или в операторной форме

$$\Delta Q(s) \approx \frac{\Delta I_6}{s}. \quad (15-296)$$

В активном режиме, когда действительно соотношение (15-256), аналогичные выражения получаются для коллекторного тока¹:

$$\Delta I_k(t) \approx \Delta I_6 \frac{t}{\tau_\alpha}; \quad (15-30a)$$

$$\Delta I_k(s) \approx \frac{\Delta I_6}{s\tau_\alpha}. \quad (15-306)$$

Выражения (15-29) и (15-30) действительны при условии постоянства производной dQ/dt . Согласно (15-24в) такое постоянство обеспечивается при выполнении неравенства

$$\Delta I_6 \gg \frac{\Delta Q}{\tau} \quad (15-31a)$$

или для активного режима [с учетом (15-256)]

$$\Delta I_6 \gg \frac{\tau_\alpha \Delta I_k}{\tau} \approx \frac{\Delta I_k}{\beta}. \quad (15-316)$$

Неравенства (15-31) называют *условиями сильного сигнала*. Значения ΔQ и ΔI_k в правых частях (15-31) легко оцениваются при конкретном анализе переходных процессов. Если неравенства (15-31) остаются в силе на протяжении всего фронта выходного импульса, то фронт оказывается почти линейным.

В тех случаях, когда нужно учитывать влияние коллекторной емкости, пользуются уравнением (15-246). Если нагрузка чисто активная (сопротивление R_k), следует подставить в (15-246) $dU_k = R_k dI_k$ и выразить dI_k с помощью соотношения (15-256). Тогда легко привести (15-246) к следующему виду:

$$\frac{dQ}{dt} + \frac{Q}{\tau_{oe}} = \frac{\tau}{\tau_{oe}} I_6,$$

¹ Чтобы получить формулы (15-30), достаточно заменить в соотношении (15-256) величины Q и I_k на ΔQ и ΔI_k , после чего приравнять правые части выражений (15-29) и соотношения (15-256). Величина τ/β принята равной τ_α . Строго говоря, вместо τ_α должно стоять время пролета (4-138) (для бездрейфовых транзисторов — это время диффузии t_D).

где постоянная времени τ_{oe} заменяется τ :

$$\tau_{oe} \approx \tau + \bar{C}_k^* R_k. \quad (15-32a)$$

Полагая $\tau = (1 + \beta) \tau_\alpha$ и деля обе части (15-32a) на $1 + \beta$, получим эквивалентную постоянную времени, которая заменяет τ_α , если нужно учесть коллекторную емкость:

$$\tau_{\alpha oe} \approx \tau_\alpha + \bar{C}_k R_k. \quad (15-32b)$$

Под емкостями $\bar{C}_k^*$ и $\bar{C}_k$ понимаются емкости, усредненные по диапазону напряжений. Согласно [147] такое усреднение приводит к результату $\bar{C}_k \approx 1,6 C_k (E_k)$ в случае отпирания, когда напряжение меняется от E_k до 0, и $\bar{C}_k \approx 2,1 C_k (E_k)$ в случае запираания, когда напряжение меняется от 0 до E_k . В сущности постоянные времени τ_{oe} и $\tau_{\alpha oe}$ те же, что и в усилительной технике [см. (7-36) и (10-7)]. Разница состоит только в том, что емкости $\bar{C}_k^*$ и C_k усреднены, а коэффициенты токораспределения γ_k^* и γ_k при постоянных времени τ и τ_α положены равными единице (в ключевых схемах редко учитывают сопротивления r_k^* и r_k).

Легко убедиться, что в изображениях (15-28) учет емкости C_k приводит к замене τ на τ_{oe} только в знаменателях и при члене $Q(0)$. Условия сильного сигнала (15-31) с учетом емкости не меняются, а в выражениях (15-30) следует заменить τ_α на $\tau_{\alpha oe}$.

15-5. ПЕРЕХОДНЫЕ ХАРАКТЕРИСТИКИ КЛЮЧА ОЭ

Процесс замыкания транзисторного ключа можно разделить на три стадии: задержка, формирование положительного фронта выходного тока и накопление избыточного заряда в базе. Процесс размыкания ключа можно разделить на две стадии: рассасывание избыточного заряда и формирование отрицательного фронта. Анализ переходного процесса в транзисторном ключе был впервые проведен в работе [155] на основе эквивалентной схемы (рис. 4-9) с использованием операторных выражений для коэффициента α . Однако проще рассмотреть каждую из пяти стадий переходного процесса, используя для анализа уравнения заряда (15-24).

Задержка фронта. Эта стадия переходного процесса обусловлена перезарядкой барьерных емкостей C_β и C_k под действием входного сигнала. В исходном состоянии, когда ключ заперт, на базе транзистора имеется положительное смещение U_{60} , обусловленное положительным входным сигналом $E_6 = E_{62}$ (рис. 15-1). Когда сигнал E_6 скачком принимает отрицательное значение $-E_{61}$, призванное отпереть ключ, через сопротивление R_6 и входную емкость $C_{вх}$ начинает протекать ток. Поскольку напряжение на входной емкости не может измениться скачком, транзистор остается запертым вплоть до полного ее разряда. Емкость $C_{вх}$ перезарядается от напряжения U_{60} до $-E_{61}$ по закону

$$u_c(t) = u_6(t) = (E_{61} + U_{60}) e^{-t/\tau_{вх}} - E_{61},$$

где $\tau_{\text{вх}} = R_6 C_{\text{вх}}$. Подставляя $U_6 = 0$, находим время разряда емкости, которое и есть время задержки фронта:

$$t_3 = \tau_{\text{вх}} \ln \left(1 + \frac{U_{60}}{E_{61}} \right). \quad (15-33)$$

Положим, что сопротивление R_k достаточно мало и поэтому емкости C_3 и C_k можно считать соединенными параллельно. Тогда $C_{\text{вх}} = C_3 + C_k$.

Пусть, например, $C_3 + C_k = 15$ пФ; $E_{61} = U_{60} = 2$ В; $R_6 = 2$ кОм; тогда $t_3 \approx 20$ нс. В ряде случаев задержка сравнима с длительностями фронтов и даже превышает их.

В другом крайнем случае ($R_k = \infty$) напряжение на емкости C_k не меняется в интервале t_3 и время задержки определяется только разрядом емкости C_3 , т. е. в формуле (15-33) нужно положить $\tau_{\text{вх}} = R_6 C_3$. В этом случае изменения потенциала U_6 будут полностью передаваться через емкость C_k на коллектор и в момент t_3 потенциал последнего будет равен $E_k + U_{60}$. Иначе говоря, последующему положительному фронту коллекторного напряжения будет предшествовать отрицательный выброс.

Поскольку задержка проявляется в сдвиге переходной характеристики и не влияет на форму фронта, в следующем разделе считается, что поступивший входной сигнал сразу отпирает транзистор.

Положительный фронт. Пусть на входе ключа в момент $t = 0$ задана ступенька тока I_{61} (рис. 15-11, а) и пусть этот ток достаточен для последующего насыщения:

$$I_{61} > \frac{E_k / R_k}{\beta}.$$

Коллекторный ток сначала нарастает так же, как в усилительном каскаде, т. е. экспоненциально, стремясь к установившемуся значению $\beta I_{61} > E_k / R_k$. Однако, достигнув величины $I_{k, \text{н}} = E_k / R_k$, ток I_k больше не может увеличиваться, и формирование фронта заканчивается (рис. 15-11, в). Подставляя I_{61} в уравнение (15-286), получаем изображение

$$Q(s) = \frac{I_{61}}{s + \frac{1}{\tau}},$$

которому соответствует оригинал

$$Q(t) = I_{61} \tau \left(1 - e^{-\frac{t}{\tau}} \right). \quad (15-34)$$

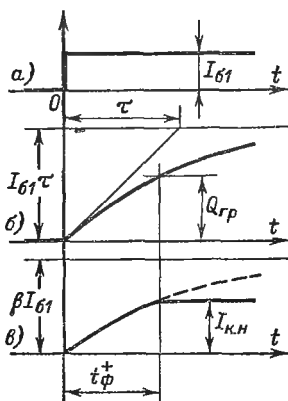


Рис. 15-11. Этап формирования положительного фронта в ключе.

а — входной ток; б — заряд в базе; в — ток в нагрузке.

Окончание положительного фронта соответствует тому моменту, когда заряд становится равным граничному значению $Q_{гр}$ (рис. 15-11, б). Подставляя (15-27) в (15-34), находим длительность положительного фронта:

$$t_{\Phi}^+ = \tau \ln \frac{I_{61}}{I_{61} - \frac{I_{к.н}}{\beta}}. \quad (15-35)$$

Например, при $\tau = 2$ мкс; $\beta = 50$; $I_{61} = 1$ мА; $I_{к.н} = 5$ мА получается $t_{\Phi}^+ = 0,2$ мкс. Если учесть коллекторную емкость, положив $C_k = 10$ пФ и $R_k = 2$ кОм, то из (15-32а) получим $\tau_{0e} \approx 3$ мкс и, следовательно, $t_{\Phi}^+ = 0,3$ мкс*.

Поскольку ток коллектора согласно (15-25) пропорционален заряду, получаем из (15-34) следующую зависимость:

$$i_k(t) = \beta I_{61} \left(1 - e^{-\frac{t}{\tau}}\right). \quad (15-36)$$

При сильном отпирающем сигнале, когда соблюдается условие (15-31б), т. е.

$$\beta I_{61} \gg I_{к.н},$$

фронт импульса близок к линейному и описывается выражением (15-30а) при $\Delta I_6 = I_{61}$:

$$i_k(t) \approx I_{61} \frac{t}{\tau_{\alpha}}. \quad (15-37)$$

Отсюда, полагая, $i_k(t_{\Phi}^+) = I_{к.н}$, легко найти длительность положительного фронта при сильном сигнале:

$$t_{\Phi}^+ \approx \tau_{\alpha} \frac{I_{к.н}}{I_{61}}. \quad (15-38)$$

Формулы (15-35) и (15-38) показывают, что длительность фронта t_{Φ}^+ уменьшается в первую очередь с ростом отпирающего тока I_{61} . При прочих равных условиях она меньше у транзисторов с меньшим временем жизни и большим значением β , в частности у дрейфовых транзисторов. При сильном сигнале величина коэффициента β не играет роли; определяющим параметром становится постоянная времени τ_{α} .

Накопление носителей. Начиная с момента t_{Φ}^+ все три внешних тока транзистора практически не меняются. Однако заряд в базе

* При наличии емкости нагрузки C_n время фронта дополнительно возрастает, поскольку емкости C_n и C_k^+ можно считать соединенными параллельно. Кроме того, как показано в [156], при условии $C_n \gg C_k$ становится заметной дополнительная задержка фронта, обусловленная зарядом емкости C_n . Зарядный ток этой емкости может намного превышать ток насыщения. Последующие выражения (15-36)—(15-38) подразумевают отсутствие емкости C_n или, во всяком случае, условие $C_n < C_k$.

продолжает нарастать по закону (15-34), и этот процесс заканчивается лишь через время

$$t_n = (2 \div 3) \tau_n, \quad (15-39)$$

которое называют *временем накопления*.

В формуле (15-39) вместо ранее использованного времени жизни τ стоит величина τ_n , которую называют постоянной времени накопления. Она отличается от величины τ , поскольку распределение носителей в базе при насыщении заметно отличается от распределения при нормальном активном режиме. Из рис. 15-12 видно, что распределение носителей при насыщении близко к распределению в инверсном активном режиме. Поэтому часто полагают [147] $\tau_n = \tau_I$, где τ_I — инверсное время жизни, которое в большей степени определяется поверхностной рекомбинацией и поэтому меньше, чем τ . Постоянная времени τ_n зависит от токов I_b и $I_{к.н}$ благодаря изменениям электронных составляющих эмиттерного и коллекторного токов.

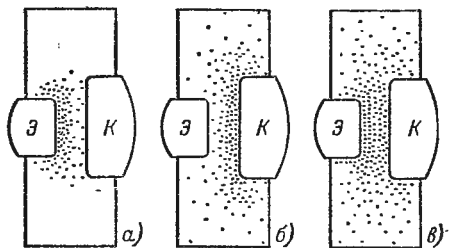


Рис. 15-12. Примерное распределение дырок в базе для нормального (а), инверсного (б) и насыщенного (в) режимов.

Рассмотрим этот вопрос на примере симметричного транзистора, для которого

$$I_n = I_{эн} + I_{кн} = I_1 (1 - \gamma) + I_2 (1 - \gamma) = (I_1 + I_2) (1 - \gamma),$$

где γ — коэффициент инжекции, а I_1 и I_2 — инжектируемые компоненты токов (см. рис. 4-9). Каждому из этих компонентов соответствует линейное распределение концентрации (рис. 15-13). Следовательно, заряды Q_1 и Q_2 связаны с токами I_1 и I_2 формулой (15-25б). Выражая сумму $I_1 + I_2$ через заряд $Q = Q_1 + Q_2$, получаем:

$$I_n = \frac{\beta (1 - \gamma)}{\tau} Q.$$

Подставляя этот ток в уравнение (15-24а) и перенося член I_n в левую часть, убеждаемся, что форма уравнения (15-24в) сохраняется, но результирующая постоянная времени имеет меньшее значение¹:

$$\tau_n = \frac{\tau}{1 + \beta (1 - \gamma)}. \quad (15-40)$$

Из (15-40) следует, что величина τ_n должна уменьшаться с ростом заряда в базе (т. е. с увеличением тока I_b), а также с ростом тока $I_{к.н}$, так как в обоих случаях уменьшается коэффициент инжекции. Кроме того, ясно, что τ_n изменяется в процессе нарастания заряда от начального значения τ до установившегося значения (15-40). Однако учет этого изменения сильно осложняет анализ; целесообразнее пользоваться усредненным значением τ_n , близкой к инверсному времени жизни, измеренному при соответствующем токе.

¹ Для несимметричного транзистора постоянная времени τ_n выражается более сложной формулой; кроме того, в правой части уравнения заряда появляется дополнительное положительное слагаемое, пропорциональное току $I_{к.н}$ и зависящее от коэффициентов инжекции.

Численный коэффициент в формуле (15-39), разумеется, условен и зависит от того уровня заряда (относительно установившегося значения $I_{61}\tau_n$), который принимается за критерий практического завершения экспоненциального процесса. Например, для уровня $0,9I_{61}\tau_n$ получится $t_n = 2,3\tau_n$ (рис. 15-14). Если отпирающий импульс не бесконечная ступенька, а имеет конечную длительность $t_{и}$, сравнимую с постоянной времени τ_n , то заряд в базе в момент $t_{и}$ будет меньше установившегося значения $I_{61}\tau_n$. Величина $Q(t_{и})$ легко находится из выражения (15-34):

$$Q(t_{и}) = I_{61}\tau_n \left(1 - e^{-\frac{t_{и}}{\tau_n}}\right). \quad (15-41)$$

Рис. 15-13. Разделение полного заряда в базе на две составляющие, соответствующие собираемому и инжектируемому компонентам коллекторного тока.

Накопление заряда в базе около переходов сопровождается увеличением положительных напряжений U_a и U_k вплоть до их установившихся значений, определяемых формулами (15-9). Рост указанных напряжений является внешним (хотя и не ярко выраженным) проявлением процесса накопления (рис. 15-14). Поскольку токи I_a и I_k в области насыщения не меняются, остаются неизменными и градиенты концентрации дырок на границах переходов. Таким образом, граничная кривая распределения (см. рис. 15-10), начиная с момента t_{ϕ}^+ , перемещается вверх параллельно самой себе. Скорость этого перемещения можно оценить, определяя производную dQ/dt из формулы (15-34) и учитывая, что $dQ \sim dp$.

В дрейфовых транзисторах, характерных сравнительно толстым и высокоомным коллекторным слоем, накопление избыточных носителей происходит не только в базе, но и в коллекторе. Действительно, считая $\rho_b \approx \rho_k$, приходим к выводу, что коэффициент инжекции коллектора в режиме насыщения далек от единицы [см. (2-35), где ρ_a нужно заменить на ρ_k]. Значит, наряду с инжекцией дырок из коллектора в базу имеет место заметная инжекция электронов из базы в коллектор. В связи с большой величиной ρ_k время жизни электронов в коллекторе сравнительно велико, а следовательно, велика и диффузионная длина. В результате в коллекторе накапливается заряд избыточных электронов и соответствующий заряд компенсирующих дырок.

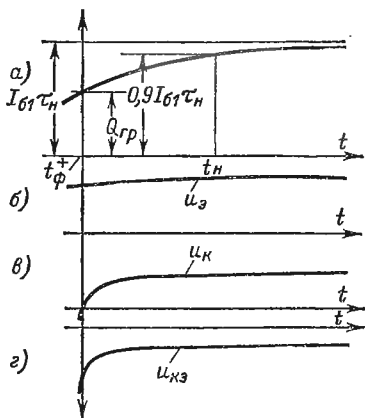


Рис. 15-14. Накопление носителей в базе.

a — заряд в базе, $б$ и $в$ — напряжения на переходах; $г$ — потенциал коллектора.

В результате в коллекторе накапливается заряд избыточных электронов и соответствующий заряд компенсирующих дырок.

В диффузионном приближении накопленный заряд можно оценить по формуле (2-84), заменяя $p(0)$ на $n(0)$ и w на w_k . Однако на самом деле в слое коллектора имеется электрическое поле, обусловленное напряжением $I_{k'rk}$, причем дрейф носителей под действием этого поля оказывается более существенным, чем диффузия их. С учетом дрейфа оценка неравновесного заряда $\Delta Q_{пк}$ согласно [157] приводит к выражению

$$\Delta Q_{пк} \approx q S_k n(0) \frac{\varphi_T w_k}{I_{k'rk}}.$$

Основная доля заряда обычно накапливается не в активной, а в пассивной области коллектора, т. е. в области, расположенной вне проекции эмиттерного перехода на коллекторный (рис. 4-36). В целом заряд коллектора может существенно превышать заряд базы.

Накопление коллекторного заряда происходит в течение того времени, когда граничная концентрация $p(w)$ повышается от нуля до установившегося значения (см. рис. 15-10), т. е. с постоянной времени τ_n .

Рассасывание носителей. Пусть в некоторый момент времени входной ток ключа скачком уменьшается от положительного значения I_{61} до отрицательного значения I_{62} (рис. 15-15, а) *.

Отрицательный перепад тока означает «отсос» заряда, и потому последний, естественно, начинает уменьшаться. Однако, как и в заряженной емкости, заряд не может измениться скачком, не может мгновенно измениться и распределение заряда в базе. Следовательно, в течение некоторого времени граничные концентрации дырок у обоих переходов остаются выше равновесной концентрации и напряжения U_b и U_k сохраняют небольшую положительную величину. Очевидно, что ток коллектора при этом практически не меняется, а ток эмиттера уменьшается скачком в начальный момент на величину $\Delta I_e = \Delta I_b$, а затем остается неизменным. Все указанные процессы объединяют названием стадии рассасывания. Главное внешнее проявление этой стадии состоит в задержке начала фронта относительно момента поступления запирающего сигнала (рис. 15-15, б). Эта задержка характеризуется *временем рассасывания* t_p — специфичным и важным параметром ключевого режима транзистора.

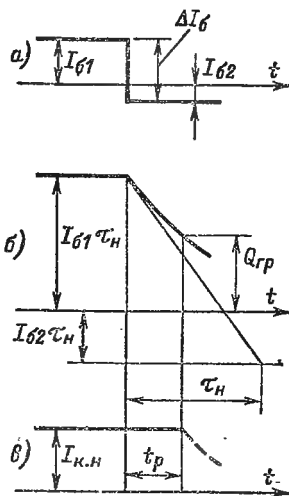


Рис. 15-15. Рассасывание носителей в базе.

а — входной ток; б — заряд в базе; в — ток нагрузки.

* Ток I_{62} может быть и положительным, но должен иметь значение $I_{62} < I_{к.н}/\beta$, достаточно малое для выхода транзистора из насыщения. Однако при положительном токе I_{62} транзистор не может запереться, т. е. вернуться к исходному состоянию. Поэтому случай $I_{62} < 0$ (рис. 15-15) более типичен.

Для анализа процесса рассасывания воспользуемся опять уравнением заряда, но в форме (15-28а), в которой учтен отличный от нуля начальный заряд. Подставив ток I_{62} и заменив τ на τ_n , получим изображение заряда¹

$$Q(s) = \frac{sQ(0) + I_{62}}{s + \frac{1}{\tau_n}},$$

которому соответствует оригинал

$$Q(t) = Q(0)e^{-\frac{t}{\tau_n}} + I_{62}\tau_n \left(1 - e^{-\frac{t}{\tau_n}}\right). \quad (15-42)$$

Рассасывание закончится тогда, когда и з б ы т о ч н ы й заряд в базе исчезнет и будет иметь место равенство $Q = Q_{гр}$. Подставив $Q_{гр}$ в левую часть (15-42), легко найти время рассасывания в следующем общем виде:

$$t_p = \tau_n \ln \frac{Q(0) - I_{62}\tau_n}{Q_{гр} - I_{62}\tau_n}. \quad (15-43а)$$

Используя выражения (15-41) и (15-27) соответственно для $Q(0)$ и $Q_{гр}$ и полагая в последнем $\tau = \tau_n$, выразим время рассасывания через управляющие токи:

$$t_p = \tau_n \ln \frac{I_{62} \left(1 - e^{-\frac{t_n}{\tau_n}}\right) - I_{62}}{\frac{I_{к.н}}{\beta} - I_{62}}. \quad (15-43б)$$

Здесь t_n — длительность отпирающего сигнала I_{61} . Если эта длительность значительно превышает величину τ_n , получается более простая формула:

$$t_p = \tau_n \ln \frac{I_{61} - I_{62}}{\frac{I_{к.н}}{\beta} - I_{62}}. \quad (15-43в)$$

Эта формула широко применяется на практике, причем нередко забывают, что ее справедливость ограничена достаточно длинным отпирающим импульсом.

Вводя в формулу (15-43в) величину запирающего импульса $\Delta I_6 = I_{61} - I_{62}$ и используя определение степени насыщения (15-8), можно привести время t_p к виду

$$t_p = -\tau_n \ln \left(1 - \frac{NI_{к.н}}{\beta \Delta I_6}\right).$$

¹ Мы не вводим специального индекса для постоянного времени рассасывания, хотя она значительно (на 20—25%) меньше, чем постоянная времени насыщения. Различие объясняется тем, что накопление начинается при малом заряде в базе, а рассасывание — при большом, когда поверхностная рекомбинация играет значительную роль. Таким образом, постоянную времени τ_n следует измерять раздельно при отпирании и запираании ключа.

В случае сильного запирающего сигнала выполняется условие¹

$$\Delta I_6 \gg \frac{NI_{к.н}}{\beta}; \quad (15-44)$$

тогда, разлагая логарифм в ряд, получаем упрощенное выражение для времени рассасывания:

$$t_p \approx \frac{\tau_n}{\beta} \frac{NI_{к.н}}{\Delta I_6}. \quad (15-45a)$$

Из этого выражения особенно ясно видно, что время рассасывания и связанная с ним задержка уменьшаются с увеличением запирающего сигнала и уменьшением степени насыщения. Последнее обстоятельство делает нежелательными большие отпирающие токи I_{61} , выгодные с точки зрения длительности положительного фронта.

Полагая $N \approx \beta I_{61}/I_{к.н}$ (что вполне справедливо при $N > 3$) и подставляя это значение N в (15-45a), получаем выражение

$$t_p \approx \tau_n \frac{I_{61}}{\Delta I_6}, \quad (15-45б)$$

из которого следует, что при сильном сигнале и достаточной степени насыщения время рассасывания практически не зависит от величины β .

Если $N = 5$; $\Delta I_6 = 2$ мА, а значения τ , β и $I_{к.н}$ те же, что и в примере к формуле (15-35), то из (15-45a) получаем $t_p = 0,5$ мкс — величину в 2,5 раза большую, чем t_{Φ}^+ . Для того чтобы сделать $t_p \approx t_{\Phi}^+$, требуется запирающий сигнал $\Delta I_6 \approx 5$ мА, т. е. ток $I_{62} = -4$ мА (если $I_{61} = 1$ мА).

Во всех предшествующих выводах подразумевалось, что отрицательный запирающий ток I_{62} достаточно мал по сравнению с коллекторным током насыщения. Только при таком условии можно считать, что в процессе рассасывания кривая распределения (см. рис. 15-10) «сползает» вниз, приблизительно сохраняя свою форму. Если же ток I_{62} близок к $I_{к.н}$ или больше его, проведенный анализ будет неточным или даже качественно неверным.

В самом деле, если $|I_{62}| > I_{к.н}$, то ток эмиттера будет в процессе рассасывания отрицательным:

$$I_9 = I_{к.н} + I_{62} < 0.$$

Это значит, что оба тока (I_9 и I_k) вытекают из базы, унося избыточный заряд. Соответственно кривая распределения концент-

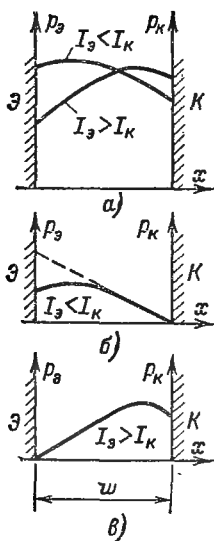
¹ В самом деле, за время рассасывания заряд в базе меняется на величину

$$\Delta Q = Q_{61} - Q_{\Gamma p} = \left(I_{61} - \frac{I_{к.н}}{\beta} \right) \tau = \frac{NI_{к.н}}{\beta} \tau.$$

Подставив это значение ΔQ в условие для сильного сигнала (15-31a), получим неравенство, приведенное в тексте.

Рис. 15-16. Распределение неосновных носителей в базе при поступлении сильного запирающего сигнала.

а — в период рассасывания; б — в конце нормального рассасывания; в — в конце инверсного рассасывания.



рации не будет монотонной, а будет иметь максимум (рис. 15-16, а). С уменьшением заряда кривая $p(x)$ опускается. При этом могут быть два основных случая в зависимости от соотношения токов $I_э$ и $I_к$, т. е. от того, у какого из двух переходов избыточная концентрация раньше падает до нуля — у коллекторного (рис. 15-16, б) или у эмиттерного (рис. 15-16, в).

Первый случай (*нормальное или коллекторное рассасывание* — рис. 15-16, б) в общем охватывается ранее приведенным анализом, хотя граничный заряд $Q_{гр}$ будет меньше, чем следует из формулы (15-27) (см. пунктирную кривую на рис. 15-16, б). Однако это отличие чисто количественное и не очень существенное, так как величина $Q_{гр}$, во-первых, стоит под логарифмом и, во-вторых, складывается с величиной $I_{бэ} \tau_n$, которая обычно значительно больше, чем $Q_{гр}$. Второй случай (*инверсное или эмиттерное рассасывание* — рис. 15-16, в) означает выход транзистора из насыщения «через эмиттер», а не «через коллектор», как считалось до сих пор¹. Это приводит к некоторым специфическим явлениям, которые будут рассмотрены ниже при анализе отрицательного фронта.

Оценим время $t_{рI}$ при инверсном рассасывании, учитывая, что после выхода из насыщения транзистор окажется в инверсном активном режиме (см. § 15-1), в котором коллектор смещен в положительном направлении, а эмиттер — в отрицательном. Следовательно, граничный заряд в данном случае можно записать по аналогии с (15-27) в следующем виде:

$$Q_{рI} = \frac{\tau_n}{\beta_I} I_э. \quad (15-46)$$

Здесь $I_э = |I_{к.н} + I_{бэ}|$ — ток эмиттера в период инверсного рассасывания и в момент выхода из насыщения. Поскольку остальные величины в формуле (15-43а) не зависят от характера рассасывания, легко получить время рассасывания $t_{рI}$, заменяя в формулах (15-43б) и (15-43в) β на β_I и ток $I_{к.н}$ на ток — $(I_{к.н} + I_{бэ})$.

¹ Инверсное рассасывание особенно характерно для дрейфовых транзисторов, поскольку у них при отрицательном эмиттерном токе дрейфовая составляющая этого тока сохраняет прежнее положительное направление, а следовательно, диффузионная (отрицательная) составляющая должна быть значительно больше. Соответственно больше должен быть положительный градиент концентрации вблизи эмиттера.

результате, например, для $I_{\text{н}} = I_{\text{н}0}$ отпирающего импульса из формулы (15-43в) найдем:

$$t_{pI} = \tau_n \ln \frac{I_{61} - I_{62}}{-\frac{I_{\text{к.н}}}{\beta_I} - \frac{I_{62}}{\alpha_I}}. \quad (15-47)$$

Очевидно, что инверсное рассасывание имеет место при выполнении неравенства $t_{pI} < t_p$, откуда с помощью выражений (15-43в) и (15-47) получаем следующее условие инверсного рассасывания:

$$|I_{62}| > I_{\text{к.н}} \left(1 + \frac{\beta_I}{\beta_N}\right). \quad (15-48)$$

Для симметричного бездрейфового транзистора (с одинаковыми площадями эмиттера и коллектора), у которого $\beta_I = \beta_N$, вместо (15-48) будем иметь условие $|I_{62}| > 2I_{\text{к.н}}$. Это вполне естественно, так как именно при этом условии получится $|I_9| > I_{\text{к.н}}$.

В предыдущих разделах, говоря о накоплении заряда в базе и граничном заряде (15-27), мы подразумевали активную область базы (см. с. 176 и рис. 4-14, где активной является область I). Между тем избыточный заряд, несомненно, накапливается и в пассивных областях базы, как видно из рис. 15-7 и 15-12, причем «пассивный» заряд, как правило, много больше «активного». С учетом «пассивного» заряда время накопления будет определяться той же формулой (15-39), но с несколько другой постоянной времени τ_n , учитывающей рекомбинацию в объеме и на поверхности пассивной части базы.

Что касается рассасывания, то оно с учетом заряда в пассивной области базы может иметь некоторую специфику [158]. А именно «активный» заряд при сильном запирающем токе довольно быстро уходит из базы через $p-n$ переходы, тогда как «пассивный» заряд частично рекомбинирует в той области, где он находится, а частично диффундирует в активную область, пополняя ее заряд. Поскольку диффузия происходит сравнительно медленно, заряд, накопленный в «далекой» пассивной области базы, может сохраниться до тех пор, когда в активной области рассасывание уже закончилось и оба перехода заперты. Последующая диффузия остаточного, «задержавшегося» заряда из пассивной области в активную может вызвать повторное отпирание транзистора и соответствующий ложный импульс в коллекторной цепи. Такое явление «самооткрывания» иногда наблюдалось у транзисторов с большим объемом базы при высокой степени насыщения и особенно при повышенной температуре, когда увеличивается время жизни.

Время рассасывания с учетом «пассивного» заряда рассчитывается по прежним формулам, но с использованием постоянной времени τ_n , свойственной пассивной области базы.

В случае дрейфовых транзисторов (см. рис. 4-36) объем активной базы весьма мал по сравнению с объемами пассивной базы и коллектора (где тоже происходит накопление носителей, см. предыдущий раздел). Поэтому у дрейфовых транзисторов обычно можно пренебречь зарядом в активной базе и считать, что время t_p соответствует рассасыванию лишь «пассивных» зарядов. Тогда анализ, выполненный в работе [157], приводит к выводу, что формулы (15-43) остаются в силе, но постоянную времени τ_n следует заменить на τ'_n . При толстом слое коллектора ($w_k > 2L_k$)

$$\tau'_n \approx \tau_n \left(1 + \frac{L_k}{w_1}\right) / \left(1 + \frac{2D_k \tau_n}{w_1 L_k}\right);$$

при тонком коллекторе ($w_k < 0,5L_k$)

$$\tau_n \approx \tau_n \left(1 + \frac{w_k}{w_1}\right) / \left(1 + \frac{2D_k \tau_n}{w_1 w_k}\right).$$

В обоих выражениях w_1 — толщина пассивной базы (см. рис. 15-7), а индексы «к» относятся к коллекторному слою.

У дрейфовых транзисторов процесс рассасывания может осложниться тем, что при достаточно большом запирающем токе I_{62} рассасывание основного заряда из активной и пассивной областей базы заканчивается весьма быстро, а «задержавшийся» заряд электронов, накопившихся в коллекторе (см. с. 482), поступая в базу, может вызвать такое же повторное «самооткрывание» транзистора, какое описано выше применительно к заряду пассивной базы.

Для того чтобы избежать этого явления, нужно обеспечить достаточную длительность запирающего импульса. В работе [157] показано, что если основная часть заряда накапливается в пассивной области коллектора, то минимальное время, необходимое для перехода этого заряда в базу, составляет: при тонком коллекторе ($w_k < 0,5L_k$)

$$t_{\text{мин}} \approx \frac{w_k^2}{5D_k},$$

а при толстом коллекторе ($w_k > 2L_k$)

$$t_{\text{мин}} \approx 2,3 \frac{L_k^2}{D_k} \approx 2,3\tau_k.$$

Как видим, структура этих выражений примерно соответствует либо времени диффузии (при тонком слое), либо времени рассасывания за счет рекомбинации (при толстом слое).

Отрицательный фронт. Рассмотрим сначала простейший случай, когда емкостями C_k и C_n можно пренебречь, а запирающий ток $|I_{62}|$ значительно меньше тока насыщения $I_{k,n}$. В этом случае заряд в базе меняется под действием тока I_{62} по тому же закону (15-42), что и на предыдущей стадии. Однако величина заряда, разумеется, не может достигать асимптотического значения $-|I_{62}| \tau$, так как заряд дырок всегда положителен. Поэтому процесс формирования отрицательного фронта заканчивается, когда $Q = 0$ (рис. 15-17, а). Таким образом, в уравнении (15-28а) начальным условием нужно считать $Q(0) = Q_{\text{гр}}$, а в полученном решении вида (15-42) следует положить $Q(t_{\text{ф}}) = 0$. Тогда время отрицательного фронта после подстановки (15-27) оказывается равным:

$$t_{\text{ф}} = \tau \ln \frac{\frac{I_{k,n}}{\beta} - I_{62}}{-I_{62}}. \quad (15-49a)$$

При сильном запирающем сигнале, когда $|I_{62}| \gg I_{k,n}/\beta$, можно разложить логарифм в ряд и получить упрощенное выражение:

$$t_{\text{ф}} \approx \frac{\tau}{\beta} \frac{I_{k,n}}{-I_{62}}; \quad (15-49б)$$

оно соответствует линейному изменению заряда. Поскольку транзистор на данной стадии работает в активном режиме, ток

коллектора изменяется пропорционально заряду и за время t_{ϕ} уменьшается от $I_k = I_{k,n}$ до $I_k = 0$ (рис. 15-17, б). Соответственно напряжение U_k возрастает от $U_k \approx 0$ до $U_k \approx E_k$ (рис. 15-17, в).

Заменяя в (15-49) параметр τ на τ_{oe} , можно учесть влияние коллекторной емкости и сопротивления нагрузки [см. (15-32а)].

В тех типичных случаях, когда запирающий ток $|I_{62}|$ сравним с током насыщения $I_{k,n}$ и больше его или когда существенна емкость нагрузки, процесс запирания ключа развивается сложнее.

Положим сначала, что емкость C_n по-прежнему равна нулю, но ток $|I_{62}|$ составляет заметную долю тока $I_{k,n}$. Тогда главная

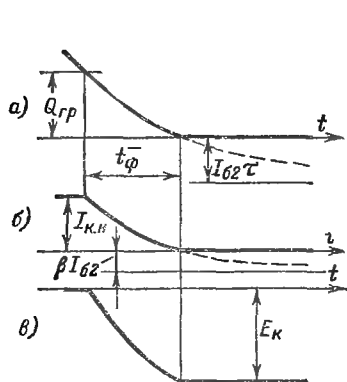


Рис. 15-17. Формирование отрицательного фронта в упрощенном виде.

а — заряд в базе; б — ток коллектора; в — коллекторное напряжение.

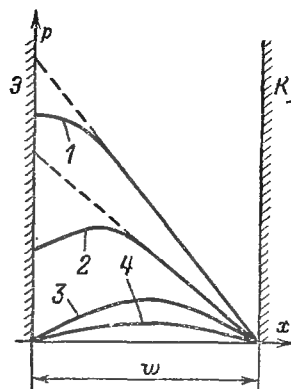


Рис. 15-18. Распределение неосновных носителей в базе при запирании транзистора.

причина неточности формул (15-49) заключается в изменениях постоянной времени и предельного значения I_{62} во время переходного процесса. Рассмотрим эту причину на примере симметричного бездрейфового транзистора, у которого в режиме насыщения распределение дырок в базе близко к линейному (рис. 15-10).

При с и л ь н о м запирающем сигнале это распределение может стать нелинейным уже на стадии рассасывания (см. рис. 15-16). Но даже при сравнительно слабом сигнале распределение неизбежно оказывается нелинейным в пределах отрицательного фронта, когда ток I_k , уменьшаясь, делается близким к $|I_{62}|$, а ток $I_{\phi} = I_k - |I_{62}|$ — близким к нулю (рис. 15-18, кривая 1). Еще большая нелинейность получается в дальнейшем, когда ток I_k становится меньше тока $|I_{62}|$, а ток I_{ϕ} делается отрицательным (рис. 15-18, кривая 2). Такого рода нелинейное распределение означает, что данному току соответствует заряд, значительно меньший, чем следует из выражений (15-25) (пунктирные линии на рис. 15-18).

Значит, начиная с тока $I_k \approx |I_{62}|$, коэффициент пропорциональности между зарядом и током в формулах (15-25) уменьшается; соответственно ток I_k изменяется медленнее, чем заряд.

Нарушение пропорциональности между током и зарядом можно приписать уменьшению «эффективной» толщины базы при отрицательных токах эмиттера, в результате чего уменьшается параметр t_D в (15-25а) и растет параметр β в формуле (15-25б). Такая интерпретация оправдана формой кривых распределения (рис. 15-18): с ростом тока $|I_3|$ протяженность квазилинейного участка между максимумом и коллектором действительно уменьшается, стремясь к $w/2$.

Когда концентрация дырок у эмиттерного перехода падает до равновесного значения (рис. 15-18, кривая 3), напряжение U_3 делается равным нулю. После этого эмиттерный переход работает с отрицательным смещением, как и коллекторный, и уменьшаются оба тока: I_k и $|I_3|$ (рис. 15-18, кривая 4). Следовательно, начиная с этого момента уменьшается и ток базы, равный их сумме.

Этап, на котором оба p - n перехода смещены в обратном направлении, но в базе (в отличие от статического режима) имеется некоторый остаточный заряд, назовем этапом *динамической отсечки* транзистора.

Этап динамической отсечки характеризуется автономностью процесса запираания: все три тока уменьшаются на этом этапе по одному и тому же закону независимо от напряжений и сопротивлений во внешних цепях электродов¹. Закон изменения — экспоненциальный:

$$i(t) = I(0) e^{-t/\tau_{\text{отс}}},$$

где $\tau_{\text{отс}}$ — *постоянная времени отсечки*, а время отсчитывается от момента, когда $U_3 = 0$. В этот начальный момент ток $I_6(0)$ еще сохраняет значение I_{62} , а токи $I_3(0)$ и $I_k(0)$ для симметричного транзистора близки к значению $1/2 |I_{62}|$, причем ток эмиттера имеет отрицательное значение (рис. 15-19, сплошные линии).

Для того чтобы оценить постоянную времени $\tau_{\text{отс}}$, представим кривые 1 и 4 на рис. 15-18 в виде равнобедренных треугольников, вершины которых расположены посередине базы. Иначе говоря, мысленно разрежем транзистор по вертикали в точке $x = w/2$ и будем считать, что в каждой половине рассасывание заряда происходит независимо, под действием уменьшающегося тока базы. Тогда время пролета носителей в каждой половине будет близко к соответствующему времени диффузии t'_D , а значит, учитывая (2-88),

$$\tau_{\text{отс}} \approx t'_D = \frac{(w/2)^2}{2D} = 0,25 t_D. \quad (15-50)$$

Полное время динамической отсечки на уровне 0,1 составляет $t_{\text{отс}} = 2,3 \tau_{\text{отс}} \approx 0,5 \tau_a$.

¹ Такая автономность свойственна также диоду на этапе восстановления обратного тока (см. § 2-9).

Таким образом, в общем случае отрицательный фронт тока состоит из трех участков: первый соответствует работе транзистора в активном режиме (постоянная времени τ_{oe}), второй является промежуточным (постоянная времени уменьшается от τ_{oe} до $\tau_{отс}$), третий соответствует режиму динамической отсечки (постоянная времени $\tau_{отс}$ и нулевой асимптотический уровень для всех трех токов). Промежуточный участок начинается при условии $I_k = |I_{62}|$ и кончается при условии $U_3 = 0$. В случае $|I_{62}| < 0,2 I_{к.н}$ превалирует 1-й участок и длительность фронта можно рассчитывать по формулам (15-49). В случае $|I_{62}| \approx I_{к.н}$, наоборот, 1-й участок практически отсутствует

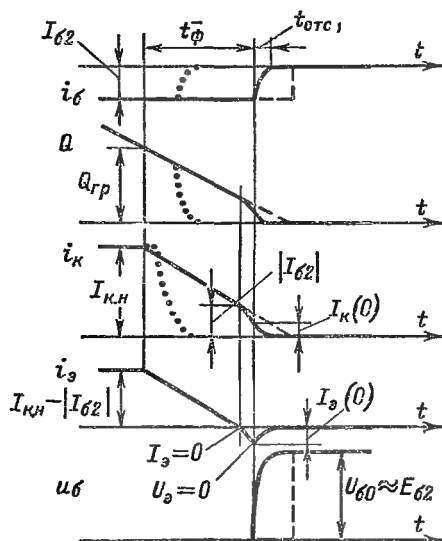


Рис. 15-19. Формирование отрицательного фронта с учетом этапа динамической отсечки.

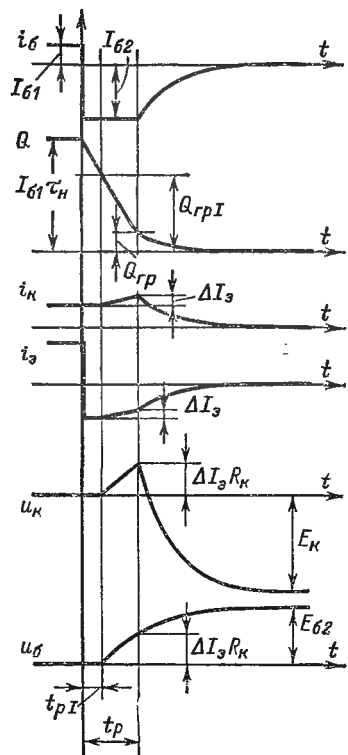


Рис. 15-20. Формирование отрицательного фронта в случае инверсного рассасывания.

и фронт формируется с постоянной времени, меньшей τ_a , т. е. практически мгновенно. Первому случаю соответствует на рис. 15-19 тонкий пунктир, а второму — жирный. При инверсном рассасывании, т. е. при условии $|I_{62}| > (1-2)I_{к.н}$, отрицательный фронт примерно такой же, как во втором случае.

Особенность инверсного режима состоит в том, что собственно отрицательному фронту предшествуют положительные выбросы коллекторного тока и коллекторного напряжения (рис. 15-20). Действительно, если эмиттерный переход выходит из насыщения раньше коллекторного, то нарастающее отрицательное

напряжение на эмиттерном переходе вызывает рост положительного потенциала базы. Приращение ΔU_6 через еще «закороченный» коллекторный переход непосредственно передается на коллектор, и потенциал последнего увеличивается [159]; соответственно увеличивается и ток коллектора. Вершина выбросов имеет место в момент выхода коллектора из насыщения (t_p), а весь отрицательный фронт формируется в условиях динамической отсечки, поскольку теперь оба тока I_3 и I_K вытекают из базы. Заметим, что данный вариант переходного процесса более характерен для дрейфовых транзисторов, у которых в режиме насыщения распределение неосновных носителей в базе более равномерное, чем на рис. 15-10 и 15-13, а при большой степени насыщения имеет нарастающий характер¹.

До сих пор длительность фронта t_f^- считалась одинаковой как для тока $i_K(t)$, так и для напряжения $u_K(t)$, поскольку при чисто активной нагрузке кривые $i_K(t)$ и $u_K(t)$ оказываются идентичными с точностью до знака. Однако если нагрузка имеет емкостно-резистивный характер (что типично для большинства реальных ключевых схем), то кривые $i_K(t)$ и $u_K(t)$ могут существенно различаться. Соответственно будут различаться и времена фронтов по току и по напряжению. Как правило, главный интерес представляет время фронта по напряжению.

Как показано в работе [156], при наличии емкости нагрузки C_H приращения коллекторного тока на 1-м участке фронта описываются операторным выражением

$$\Delta I_K(s) = \beta \Delta I_6 \frac{1 + s\tau_H}{(1 + s\tau_{\text{ЭКВ}})(1 + s\tau_0)},$$

где постоянные времени имеют значения:

$$\tau_H = C_H R_K;$$

$$\tau_{\text{ЭКВ}} = (\tau_{oe} + \tau_H);$$

$$\tau_0 \approx \frac{\tau}{\tau_{\text{ЭКВ}}} (C_K + C_H) R_K.$$

Учитывая неравенство $\tau_0 \ll \tau_{\text{ЭКВ}}$, можно считать параметр τ_0 временем задержки (см. сноску на с. 199). Тогда

$$\Delta I_K(s) \approx \beta \Delta I_6 e^{-s\tau_0} \frac{1 + s\tau_H}{1 + s\tau_{\text{ЭКВ}}}.$$

Из этого выражения следует, что без учета задержки функция $\Delta I_K(t)$ в начальный момент времени ($s = \infty$) характеризуется скачком тока, равным:

$$\Delta I_K(0) = \beta \Delta I_6 \frac{\tau_H}{\tau_{\text{ЭКВ}}}.$$

Фактически этот «скачок» имеет длительность, близкую к τ_0 , т. е. значительно меньшую, чем следует из (15-49а). Коллекторный ток в конце «скачка» имеет величину $I_{K.H} \sim |\Delta I_K(0)|$. Если эта величина меньше, чем $|I_{62}|$, то можно считать, что весь 1-й участок фронта формируется с малой постоянной

¹ Такой вывод следует из кривых на рис. 4-30 и 4-42, если мысленно их сложить (что примерно соответствует условиям в режиме двойной инжекции).

времени τ_0 . Поскольку длительность остальных двух участков всегда мала (см. выше), приходим к выводу, что при условии

$$I_{K.н} - |\Delta I_K(0)| \leq |I_{62}|$$

спад коллекторного тока имеет почти ступенчатый характер (см. рис. 15-19). Используя значение $\Delta I_K(0)$ и полагая $\beta |I_{62}| \gg I_{K.н}$, запишем это условие в следующем виде:

$$\tau_n \geq \tau_{oe} \frac{I_{K.н} - |I_{62}|}{\beta |I_{62}|}. \quad (15-51)$$

Например, если $|I_{62}| = 0,5 I_{K.н}$, то фронт тока можно считать ступенчатым при $\tau_n \geq \tau_{oe}/\beta \approx \tau_{\alpha oe}$. У дрейфовых транзисторов эквивалентная постоянная времени (15-22б), как правило, определяется вторым слагаемым; тогда предыдущее неравенство принимает вид: $C_n \geq C_K$.

В большинстве реальных транзисторных ключей условие (15-51) выполняется даже в случае бездрейфовых транзисторов. Поэтому при анализе импульсных схем спад тока обычно считается мгновенным, однако в некоторых случаях (малая емкость нагрузки или слабый запирающий ток базы) такое допущение может оказаться неприемлемым.

В отсутствие коллекторного тока ключ превращается в пассивную RC -схему, в которой восстановление коллекторного напряжения происходит с постоянной времени $\tau_{н.к} = (C_n + C_K) R_K$ (влиянием емкости эмиттерного перехода пренебрегаем). Тем самым длительность отрицательного фронта напряжения (при ступенчатом спаде тока) составляет:

$$t_{\bar{\phi}} \approx 2,3 \tau_{н.к}. \quad (15-52)$$

Ключ с форсирующей емкостью. Назначение форсирующей емкости (рис. 15-21) заключается в том, чтобы временно (на время переходных процессов) увеличивать токи базы I_{61} и I_{62} по сравнению с их значениями, определяемыми сопротивлением R_6 . Так, при отпирании ключа вместо тока

$$I_{61} = \frac{E_{61}}{R_T + R_6 + r_6} \approx \frac{E_6}{R_6}$$

получится в первый момент гораздо больший ток

$$I_{61}(0) \approx \frac{E_{61}}{R_T + r_6}.$$

Конечно, по мере заряда емкости C ток I_{61} будет уменьшаться, но все же положительный фронт формируется током, близким к максимальному значению $I_{61}(0)$, т. е. весьма быстро. В то же время установившийся заряд в базе будет соответствовать гораздо мень-

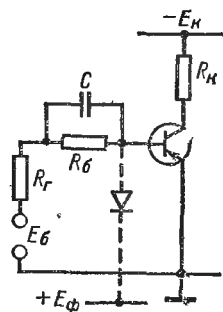


Рис. 15-21. Ключ ОЭ с форсирующей емкостью (пунктиром показана диодная фиксация базового потенциала).

шему току $I_{61}(\infty) = I_{61}$. Таким образом, форсирующая емкость позволяет получить крутой положительный фронт при слабом последующем насыщении.

Аналогично обстоит дело при запираании транзистора, когда в начальный момент вместо тока

$$I_{62} = \frac{E_{62}}{R_r + R_6 + r_6} \approx \frac{E_{62}}{R_6}$$

действует гораздо больший ток

$$I_{62}(0) \approx \frac{E_{62} + U_c(0)}{R_r + r_6} \approx \frac{E_{62} + E_{61}}{R_r + r_6};$$

это способствует более быстрому рассасыванию избыточного заряда¹. Если, однако, больший ток I_{62} сохранится до конца интервала t_p , то может иметь место и н е р с н о е рассасывание, которое часто нежелательно в связи с выбросами (см. рис. 15-20).

Величина форсирующей емкости выбирается из следующих соображений. Пусть задана отпирающая э. д. с. E_{61} в виде ступенчатого сигнала. Тогда изображение базового тока будет иметь вид:

$$I_{61}(s) = I_{61} \frac{1 + s\tau_c}{1 + s\tau'_c},$$

где

$$I_{61} = \frac{E_{61}}{R_r + R_6 + r_6};$$

$$\tau_c = CR_6;$$

$$\tau'_c = C[R_6 \parallel (R_r + r_6)] \ll \tau_c.$$

Последнее неравенство объясняется тем, что $R_r + r_6 \ll R_6$. Подставляя $I_{61}(s)$ в формулу (15-28б), получаем изображение заряда:

$$Q(s) = I_{61}\tau \frac{1 + s\tau_c}{(1 + s\tau'_c)(1 + s\tau)}.$$

Оригинал такой функции при $\tau_c > \tau$ имеет одиночный (апериодический) выброс, что, однако, не дает существенного сокращения времени фронта. Поэтому постоянную времени τ_c выбирают из условия «критического режима» (без выброса), для которого $\tau_c = \tau$ и, следовательно,

$$C = \frac{\tau}{R_6}. \quad (15-53)$$

Например, если $\tau = 2$ мкс и $R_6 = 5$ кОм, то $C = 400$ пФ.

¹ Следует иметь в виду, что сопротивление r_6 при запираании транзистора меньше, чем при отпирании, из-за модуляции сопротивления базы носителями, накопленными в режиме насыщения.

Условие $\tau_c = \tau$ приводит к оригиналу, имеющему вид простейшей экспоненциальной функции с постоянной времени $\tau'_c \ll \tau$. Соответственно время t_{Φ}^* будет значительно меньше, чем в отсутствие емкости, так как в формуле (15-35) параметр τ окажется замененным на τ'_c . По той же причине уменьшится время накопления (15-39). В то же время степень насыщения, а значит, и избыточный заряд останутся такими же, как без емкости, поскольку они определяются с т а ц и о н а р н ы м током базы.

При подаче з а п и р а ю щ е й ступеньки E_{62} процессы рассасывания и формирования заднего фронта тоже протекают форсированно, с постоянной времени τ'_c . При этом, как можно показать, условие $\tau_c = \tau$ означает, что заряд на емкости в каждый данный момент равен заряду в базе, т. е.

$$Cu_c(t) \approx Q(t).$$

Следовательно, в момент начала динамической отсечки, когда в базе еще имеется остаточный заряд, напряжение на конденсаторе имеет конечное значение, пропорциональное этому заряду.

Анализ показывает, что при $R_r + r_6 \ll R_6$ остаточное напряжение близко к величине

$$U_{\text{ост}} \approx \frac{E_6}{\beta \frac{R_r + r_6}{R_6} - 1}. \quad (15-54)$$

Поскольку во время динамической отсечки ток базы очень быстро падает до нуля, он не успевает разрядить емкость. Поэтому напряжение $U_{\text{ост}}$ уменьшается лишь благодаря разряду конденсатора C через резистор R_6 , т. е. с большой постоянной времени CR_6 . Таким образом, после запираания транзистора на его базе имеет место *динамическое* (временное) *смещение*, которое накладывается на с т а т и ч е с к о е смещение $U_{60} \approx E_{62}$.

Наличие динамического смещения приводит к некоторым осложнениям в работе ключа: суммарное напряжение U_6 оказывается больше, чем в статическом режиме и, если очередной отпирающий импульс поступает до спада динамического смещения, то задержка и длительность положительного фронта будут больше, чем следует из формул (15-33) и (15-35).

Для того чтобы избежать указанных осложнений, часто используют *диодную фиксацию* базового потенциала на уровне $E_{\Phi} = E_{62}$ или меньшем, вплоть до нуля (рис. 15-21, показано пунктиром). В этом случае, если идеализировать диод (т. е. пренебречь прямым падением напряжения на нем), потенциал базы не может стать выше уровня фиксации E_{Φ} . В случае дрейфовых транзисторов, характерных малым напряжением пробоя эмиттерного перехода, диодная

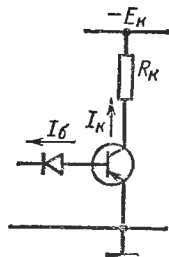


Рис. 15-22. Включение защитного диода в цепь базы дрейфового транзистора.

фиксация может предотвратить пробой, если $E_{\phi} < U_{\text{проб.}}$ Однако поскольку главная опасность пробоя состоит в большом токе базы, использование дрейфовых транзисторов, независимо от наличия или отсутствия диодной фиксации, часто сопровождается включением так называемых защитных диодов (рис. 15-22), обратный ток которых ограничивает ток базы.

15-6. РАЗНОВИДНОСТИ НАСЫЩЕННЫХ КЛЮЧЕЙ

Помимо подробно изученного ключа ОЭ иногда встречаются, главным образом в качестве элементов импульсных схем, ключи ОБ, ОК и ключ-звезда.

К л ю ч О Б (рис. 15-23, а) невыгоден тем, что входной управляющий ток должен превышать ток нагрузки, поскольку условие насыщения имеет вид:

$$\alpha_N I_{\text{э}} > I_{\text{к.}}$$

Кроме того, напряжение $U_{\text{кб}}$ на открытом ключе больше, чем напряжение $U_{\text{кэ}}$ в схеме ОЭ. Единственным, но не всегда существенным преимуществом ключа ОБ является минимальный входной ток в режиме отсечки [см. (15-1а)].

К достоинствам к л ю ч а О К (рис. 15-23, б) следует отнести минимальный ток в запертом состоянии ключа и минимальное напряжение на ключе в открытом, насыщенном состоянии (см. § 15-3). Существенный недостаток ключа ОК состоит в том,

что управляющее напряжение $E_{\text{б}}$ должно превышать напряжение питания. В самом деле, данная схема имеет структуру эмиттерного повторителя; значит, для того чтобы изменить выходное напряжение $U_{\text{э}}$ от нуля до $E_{\text{к}}$ нужно изменить входное напряжение на такую же и даже несколько большую величину.

К л ю ч - з в е з д у (рис. 15-23, в) можно рассматривать как своеобразную комбинацию ключей ОЭ и ОК, поскольку сопротивления $R_{\text{к}}$ и $R_{\text{э}}$ включены как в коллекторную, так и в эмиттерную цепь транзистора. Условие запираения, как и в ключе ОЭ, требует положительной величины $E_{\text{б}}$. Условие насыщения можно получить, принимая насыщенный транзистор за эквипотенциальную точку, выражая потенциал этой точки $U_{\text{Т}}$ по формуле (15-11в) и подставляя токи $I_{\text{к}} = (E_{\text{к}} - U_{\text{Т}}) g_{\text{к}}$ и $I_{\text{б}} = (E_{\text{б}} - U_{\text{Т}}) g_{\text{б}}$ в (15-7). Тогда

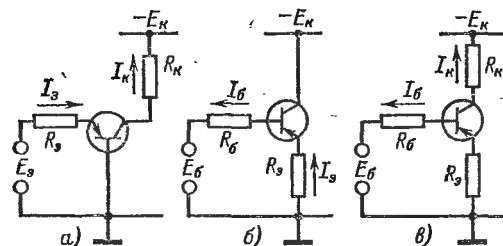


Рис. 15-23. Разновидности транзисторных ключей.

а — ключ ОБ; б — ключ ОК; в — ключ-звезда.

после преобразований можно записать критерий насыщения в таком виде:

$$E_6 > E_k \frac{(1+\beta) \frac{R_3}{R_6} + 1}{(1+\beta) \frac{R_3}{R_6} + \beta \frac{R_k}{R_6}}, \quad (15-55)$$

где E_6 и E_k — модули соответствующих напряжений. В частности, при $R_3 = 0$ этот критерий переходит в условие (15-7) для ключа ОЭ.

Интересно отметить, что при условии

$$E_6 \frac{R_3}{R_3 + R_6} > E_k \quad (15-56)$$

потенциал транзистора U_T оказывается больше, чем E_k , и, следовательно, ток коллектора становится отрицательным: транзистор вырождается в двойной диод.

Переходные процессы в ключе-звезде имеют некоторую специфику. Так, положительный фронт формируется в условиях меняющегося тока базы:

$$i_6 = I_6(0) - \gamma_6 i_k,$$

где $I_6(0) = E_{61}/(R_3 + R_6)$ — начальный ток базы; $\gamma_6 = R_3/(R_3 + R_6)$ — коэффициент обратной связи, хорошо известный по усилительным схемам.

Подставляя изображение $I_6(s) = I_6(0) - \gamma_6 I_k(s)$ в (15-28б) и выражая ток $I_k(s)$ через заряд $Q(s)$ согласно (15-25б), нетрудно получить изображение заряда. Затем, переходя к оригиналу $Q(t)$ и полагая $Q(t_\Phi^+) = Q_{гр}$, находим время положительного фронта:

$$t_\Phi^+ = \frac{\tau}{1 - \beta \gamma_6} \ln \frac{I_6(0)}{I_6(0) - \frac{I_{к.н}}{\beta} (1 + \beta \gamma_6)}. \quad (15-57)$$

Выражение (15-57) имеет две особенности: во-первых, постоянная времени перед логарифмом меньше, чем τ , и зависит от схемного параметра γ_6 , во-вторых, ток насыщения $I_{к.н}$ является функцией отпирающего сигнала E_6 , поскольку $I_{к.н} = (E_k - U_T) g_k$, где потенциал U_T зависит от E_6 . В частности, если соблюдается условие (15-56), то $I_{к.н} < 0$. При этом из (15-57) следует: $t_\Phi^+ < 0$, т. е. положительный фронт формируется мгновенно. Физически это понятно: в таком режиме транзистор работает как двойной диод, т. е. инжекция через коллекторный переход наступает одновременно с инжекцией через эмиттерный переход; следовательно, накопление избыточных носителей начинается сразу после подачи сигнала.

Процесс накопления происходит, как и в любом ключе, с постоянной времени τ_n , а время накопления определяется формулой (15-39).

Процесс рассасывания имеет ту специфику, что поступление запирающего импульса ΔE_6 сопровождается изменением не только базового тока, но и коллекторного тока насыщения. Это следует из пропорциональности величин $\Delta I_k \sim \Delta U_T \sim \Delta E_6$. Легко убедиться, что запирающий сигнал ΔE_6 вызывает увеличение тока коллектора и соответственно граничного заряда $Q_{гр}$. Поэтому при тех же значениях I_{62} и $Q(0)$ рассасывание в ключе-звезде происходит быстрее, чем в ключе ОЭ.

Формирование отрицательного фронта происходит в условиях меняющегося тока базы. Тем же методом, который использован при

выводе (15-57), можно получить:

$$t_{\Phi} \approx \frac{\tau}{1 + \beta \gamma_6} \ln \frac{-I_{62} + \frac{I'_{к.н}}{\beta} (1 + \beta \gamma_6)}{-I_{62}}, \quad (15-58)$$

где $I'_{к.н}$ — ток насыщения после подачи запирающего сигнала.

Формула (15-58) действительна только при достаточно слабом сигнале, когда динамическая отсечка играет сравнительно малую роль. В противном случае время t_{Φ} будет меньше, так как величину $I'_{к.н}$ нужно заменить на $I'_{к.н} - I_6(t_{\Phi})$.

15-7. НЕНАСЫЩЕННЫЕ КЛЮЧИ

Задержка отрицательного фронта, связанная с рассасыванием избыточного заряда, затрудняет работу многих импульсных схем. Поэтому в свое время встала задача ликвидации такой задержки путем предотвращения насыщенного режима ключа.

Казалось бы, самый простой путь — это ограничить величину отпирающего сигнала (тока базы) на таком уровне, который недостаточен для насыщения. Однако практически, с учетом разброса параметров и других факторов, этот путь реализуется только в не-

которых частных схемах (см. ниже «Токовые ключи»).

Попытки использовать диодную фиксацию коллекторного потенциала на отрицательном уровне не дали желаемых результатов: несмотря на отсутствие насыщения, задержка отрицательного фронта остается из-за рассасывания заряда в диоде и даже может превышать задержку в насыщенном ключе [147]. Ана-

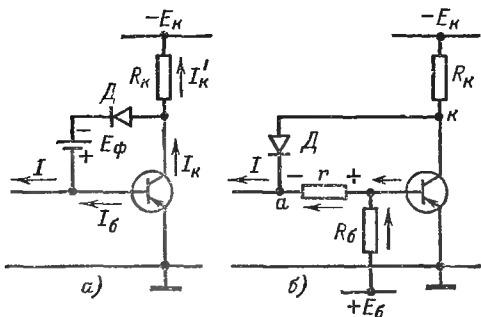


Рис. 15-24. Ключ с нелинейной обратной связью.

а — простейший вариант; б — рабочая схема.

лиз показал, что правильное решение состоит в фиксации коллекторного потенциала не на постоянном уровне, а относительно потенциала управляющего электрода. Такой способ получил название *нелинейной обратной связи* [147].

Ключи с нелинейной обратной связью. Простейший вариант такой схемы показан на рис. 15-24, а. Здесь источник смещения E_{ϕ} играет вспомогательную роль: он призван скомпенсировать прямое падение напряжения на открытом диоде. Эта задача рассматривается ниже. При анализе схемы целесообразно считать источник смещения отсутствующим, а диод идеальным (прямое напряжение и прямое сопротивление равны нулю).

При достаточно больших коллекторных напряжениях диод заперт и цепь обратной связи отключена. Отпирание диода проис-

ходит тогда, когда с увеличением тока потенциал коллектора падает до напряжения U_6 . Это условие можно записать в виде

$$E_k - I'_k R_k = U_6,$$

где I'_k — ток, соответствующий отпиранию диода¹.

Коллекторному току I'_k соответствует ток базы I'_6 [он легко определяется из соотношения (4-72)] и равный ему входной ток I' :

$$I' = I'_6 = \frac{E_k - U_6}{\beta R_k} - I_{k0}. \quad (15-59)$$

Поскольку диод принят идеальным, то после его отпирания напряжение $U_{кб}$ оказывается равным нулю независимо от дальнейшего увеличения входного тока. Следовательно, инжекции через коллекторный переход не происходит, режим насыщения не имеет места и транзистор по-прежнему работает в активном режиме.

После отпирания диода последующие приращения входного тока складываются из приращений базового тока и тока диода:

$$\Delta I = \Delta I_6 + I_d,$$

где $\Delta I = I - I'$ и $\Delta I_6 = I_6 - I'_6$. Поскольку при открытой цепи обратной связи приращения ΔI_k идут практически полностью через низкоомную цепь диода (а не через сравнительно высокоомную цепь R_k), можно считать $I_d = \Delta I_k = \beta \Delta I_6$. Тогда после некоторых преобразований получаем:

$$\Delta I = I_d / \alpha.$$

В частности, при входном токе I_1 получаем соотношение

$$I_1 - I' = I_d / \alpha. \quad (15-60)$$

Теперь подадим на вход запирающий импульс ΔI . Этот импульс почти полностью пойдет в базу транзистора, так как сопротивление участка база — эмиттер гораздо меньше сопротивления участка «диод — R_k » (в который, вообще говоря, могла бы ответить часть тока ΔI). Под действием запирающего импульса ток коллектора начинает уменьшаться. Поскольку диод остается пока открытым, приращения ΔI_k по-прежнему идут через диод в цепь базы и накладываются на исходный импульс ΔI , поэтому ток базы постепенно начинает уменьшаться. В операторном виде можно записать:

$$\Delta I_6(s) = \Delta I - \beta(s) \Delta I_6(s);$$

$$\Delta I_d(s) = \beta(s) \Delta I_6(s).$$

¹ Ток I'_k является аналогом тока $I_{к.н}$ в насыщенном ключе: в обоих случаях — это ток в нагрузке при открытом состоянии ключа. Количественное различие в значениях I'_k и $I_{к.н}$ обусловлено различием напряжений U_6 и $U_{кэ}$ [см. (15-56)]. Первое составляет 0,5—1,0 В, а второе 0,1—0,5 В (в зависимости от типа транзистора).

Отсюда нетрудно получить изображение

$$\Delta I_d(s) = \alpha(s) \Delta I.$$

Вычитая оригинал этого изображения из начального значения $I_{д1}$ [см. (15-60)], получаем на участке запираания функцию

$$i_d(t) = \alpha[(I_1 - I') - \Delta I(1 - e^{-t/\tau_a})]. \quad (15-61)$$

Время задержки отрицательного фронта t_s соответствует моменту запираания диода, когда цепь обратной связи отключается и выходное напряжение начинает возрастать до величины E_k . Полагая $i_d(t_s) = 0$, из (15-61) получаем:

$$t_s = \tau_a \ln \frac{\Delta I}{\Delta I - (I_1 - I')}. \quad (15-62)$$

Легко убедиться, что при условии $\Delta I > 1,6(I_1 - I')$ время задержки будет меньше τ_a , т. е. практически им можно пренебречь.

Физические причины резкого уменьшения задержки в ключе с нелинейной обратной связью по сравнению с насыщенным ключом сводятся к следующему. Граничный заряд $Q_{гр}$ в обоих случаях практически одинаков, так как различие между токами $I_{кн}$ и I'_k обычно несущественно. Однако начальный заряд $Q(0)$ при одном и том же входном токе значительно меньше в ключе с нелинейной обратной связью, поскольку ток базы меньше входного тока. Кроме того, в ключе с нелинейной обратной связью запирающий входной ток базы оказывается больше, чем в насыщенном ключе.

К сожалению, ненасыщенный ключ имеет и отрицательные свойства, которые обусловлены работой транзистора в активном режиме. Так, напряжение $U_{кз}$ на открытом ключе больше, чем в насыщенном режиме, и составляет 0,5 В и более. Кроме того, входное сопротивление ненасыщенного ключа сравнительно велико: оно определяется формулой (7-2). Поэтому, в частности, скачки напряжения ΔE_k проходят через диод на базу и дают кратковременные «пики» на базе и коллекторе; величина этих «пиков» определяется соотношением сопротивлений $R_{вх}$ и R_k , образующих плечи делителя напряжения. В результате ненасыщенный ключ оказывается более чувствительным к наводкам в цепи питания, т. е. обладает меньшей помехоустойчивостью.

Если диод не идеальный (как считалось до сих пор), т. е. если $U_d \neq 0$, то в открытом состоянии ключа на коллекторном переходе будет прямое смещение, равное U_d . Значит, транзистор будет находиться в режиме насыщения. Чтобы этого избежать, включают, как уже отмечалось, источник смещения $E_\phi \approx U_d$ (рис. 15-24, а). Тогда схема работает так же, как описано выше, за исключением того, что цепь обратной связи включается не сразу и не при $U_k = U_6$, а постепенно (в соответствии с ВАХ диода), начиная с $U_k = E_\phi$.

Конечно, схема на рис. 15-24, а не является рабочей, так как использование несаземленного источника E_ϕ , тем более в виде гальванической батареи, практически неприемлемо. В реальных случаях смещение E_ϕ получают

за счет падения напряжения в цепи обратной связи от того или иного тока. Пример такой схемы показан на рис. 15-24, б. В отсутствие входного тока потенциал базы близок к $+E_6$ и транзистор заперт. Отпирание его происходит при входном токе $I = E_6 / R_6$, когда $U_6 \approx 0$.

Цепь обратной связи вступает в действие при $U_K = U_a$, т. е. когда

$$E_K - I'_K R_K = \left(\frac{E_6}{R_6} + \frac{I'_K}{\beta} \right) r + U_a.$$

Отсюда легко получить ток I'_K , а затем ток I'_6 и смещение $E_\Phi = \left(I'_6 + \frac{E_6}{R_6} \right) r$.

Сопротивление r целесообразно выбирать таким, чтобы значение E_Φ несколько превышало максимальное падение напряжения на диоде при входном токе I_1 . Такой расчет приводит к условию

$$r \geq \frac{U_{д. макс.}}{\frac{E_K}{\beta R_K} + \frac{E_6}{R_6}}. \quad (15-63)$$

В тех схемах, где не требуется начальное запирающее напряжение транзистора, цепь E_6, R_6 может быть устранена.

В заключение заметим, что истинная задержка в ненасыщенном ключе соответствует формуле (15-62) только в том случае, когда времена рассасывания и восстановления диода (см. § 2-9) ничтожно малы. В противном случае задержка обычно определяется инерционностью диода, даже при условии $\Delta I \gg I_1 - I'$ [см. (15-62)]. До последнего времени это обстоятельство ограничивало возможности ненасыщенных ключей в наносекундном диапазоне.

В настоящее время в качестве диодов, включаемых в цепь обратной связи ключа (особенно в интегральных схемах), используются диоды Шоттки (см. § 3-4), у которых отсутствует накопление, а значит, и рассасывание неосновных носителей в базе. Поэтому ненасыщенные ключи с диодами Шоттки полностью описываются приведенными выше формулами и обладают максимальным быстродействием. Ключи с диодами Шоттки в цепи нелинейной обратной связи называют *транзисторами Шоттки*. В случае кремниевых транзисторов относительно малые прямые напряжения диодов Шоттки (0,3—0,5 В по сравнению с 0,7—0,8 В на переходах) позволяют исключить источник смещения E_Φ , так как при прямом напряжении 0,3—0,5 В на коллекторном переходе инжекция практически отсутствует и насыщения не наступает.

Токовые ключи. Термин «токовый ключ» (на наш взгляд, неудачный) относится к одному из важнейших типов как называемых *логическими элементами* (или *логическими вентилями*), составляющих основу современных цифровых устройств, прежде всего ЭВМ. Все типы логических элементов [160, 161] представляют собой совокупность транзисторных ключей, выполняющих ту или иную логическую функцию. Используя классификацию логических элементов по «типам логики», будем называть токовые ключи схемами ТЛЭС (транзисторная логика с эмиттерной связью). Из всех типов логик ТЛЭС — единственная, в которой используется ненасыщенный режим транзисторов, обеспечиваемый путем простого огра-

ничения уровней сигнала, без обратной связи. Именно поэтому как пример параметров (а не схемотехнически) ненасыщенных ключей мы рассматриваем схемы ТЛЭС в данном параграфе.

Простейшая схема ТЛЭС показана на рис. 15-25. Ее характерным элементом является общий резистор R_0 , к которому присоединены эмиттеры всех транзисторов (на рис. 15-25 количество тран-

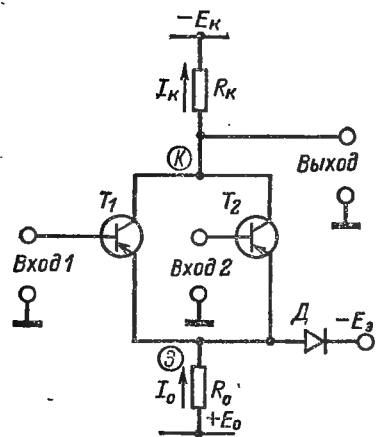


Рис. 15-25. Простейшая схема ТЛЭС.

зисторов для простоты ограничено двумя). Коллекторный резистор R_k также общий, но это не специфично для ТЛЭС и имеет место в других типах логик. Общая эмиттерная точка связана через диод D с источником смещения $-E_3$, потенциал которого меньше (по модулю) потенциала источника питания $-E_k$. Резистор R_0 питается от особого источника питания $+E_0$ положительной полярности. Наличие этого источника не обязательно, но, как увидим, облегчает построение схемы. В цепях баз отсутствуют резисторы, т. е. ключи T_1 и T_2 управляются напряжениями ¹.

Схема работает следующим образом ². В начальном состоянии транзисторы заперты и ток I_0 протекает через диод D к зажиму $-E_3$. Значение тока легко оценить из выражения

$$I_0 = \frac{E_0 - U_3}{R_0}, \quad (15-64)$$

где в начальном состоянии

$$U_3 = -E_3 + U_d < 0 \quad (15-65)$$

(U_d — прямое падение напряжения на диоде).

Для того чтобы транзисторы были действительно заперты, нужно на входы подать достаточно малый (по модулю) сигнал $U_{„0”\text{вх}}$, удовлетворяющий условию $|U_{„0”\text{вх}}| < |U_3|$. Тогда на выходе схемы потенциал будет иметь максимальное (по модулю) значение $U_{„1”\text{вых}}$, близкое к E_k .

¹ Конечно, при необходимости можно внутреннее сопротивление источника сигнала и собственное сопротивление базы рассматривать как эквивалент внешнего резистора и пересчитывать входное напряжение в ток базы.

² В дальнейшем применяются индексы «0» и «1», которые по терминологии логических элементов означают соответственно «уровень логического нуля» и «уровень логической единицы» на входе или на выходе. Учитывая отрицательную полярность рабочих сигналов в рассматриваемой схеме, будем оперировать модулями и уровнями, т. е. считать уровень «1» больше уровня «0».

Если теперь на один из входов подать достаточно большой сигнал $U_{„1”\text{вх}}$, удовлетворяющий условию $|U_{„1”\text{вх}}| > E_3$, то соответствующий транзистор откроется и потенциал в общей эмиттерной точке будет иметь величину

$$U'_3 = -U_{„1”\text{вх}} + U_{36}. \quad (15-66)$$

При этом напряжение на диоде составляет $(U'_3 + E_3) < U_{36}$, т. е. будет в зависимости от силы неравенства $|U_{„1”\text{вх}}| > E_3$ либо отрицательным, либо небольшим положительным, но недостаточным для протекания сколько-нибудь заметного тока. Таким образом, при открытом транзисторе диод D может считаться запертым. Поэтому ток I_0 вместо диода пойдет через открытый транзистор¹ и далее через резистор R_K . Потенциал на выходе уменьшится до величины $U_K = E_K - I_0 R_K = |U_{„0”\text{вых}}|$.

Для того чтобы избежать насыщения, нужно выполнить неравенство $|U_K| > |U_6|$, т. е. выбрать сопротивление R_K из условия

$$|U_{„0”\text{вых}}| = E_K - I_0 R_K > |U_{„1”\text{вх}}|. \quad (15-67)$$

Если сигнал $U_{„1”\text{вх}}$ подается на оба входа, режим схемы почти не отличается от предыдущего, так как напряжение U_{36} слабо зависит от тока, а значит, выражение (15-66) практически не изменится. В данном случае ток I_0 поделится между обоими открытыми транзисторами², а выходной потенциал $U_{„0”\text{вых}}$ останется таким же, как и при одном открытом транзисторе. Этот результат — независимость выходного тока и выходного напряжения от числа открытых транзисторов — обусловлен постоянством тока I_0 , т. е. наличием токозадающей цепи $E_0 - R_0 - D - E_3$. В отсутствие такой цепи (при заземленных эмиттерах) отпирание каждого транзистора сопровождалось бы дополнительными приращениями тока I_K и выходного потенциала U_K , т. е. непостоянством уровня $U_{„0”\text{вых}}$.

Логическая функция, выражающаяся в появлении логического нуля на выходе при любой комбинации логических единиц на входе, называется функцией ИЛИ—НЕ. Именно эту функцию выполняет схема ТЛЭС.

Поскольку насыщение транзисторов и последующее рассасывание предотвращены условием (15-67), схема ТЛЭС (как при отпирании, так и при запираании транзисторов) работает в режиме обычного усилительного каскада ОЭ, а значит, ее инерционность характеризуется постоянной времени (7-41). Учитывая весьма боль-

¹ Строго говоря, при отпирании транзистора ток I_0 несколько меняется, так как $|U'_3| > |U_3|$. Однако если неравенство $|U_{„1”\text{вх}}| > E_3$ не очень сильное и если $E_0 \gg |U_3|$ (оба условия выполняются на практике), то изменением тока I_0 можно пренебречь. Кстати, в отсутствие источника питания E_0 неизменность тока I_0 можно обеспечить только путем существенного повышения напряжений E_K и E_3 .

² Собственно из такого перераспределения тока I_0 и происходят названия «токовые переключатели» или «токовые ключи»: один и тот же исходный ток I_0 «переключается» из цепи $E_3 - D$ через один или несколько транзисторов на выход схемы.

шое значение сопротивления R_0 в цепи эмиттера, можно считать $\gamma_6 = 1$. Тогда согласно (7-41б) для одного транзистора

$$\tau_b = \tau_a + C_k R_k. \quad (15-68a)$$

Если одновременно переключаются n транзисторов, их можно считать включенными параллельно и вместо (15-68a) получится более общее выражение:

$$\tau_b = \tau_a + n C_k R_k. \quad (15-68б)$$

С физической точки зрения увеличение второго слагаемого объясняется тем, что один и тот же ток I_0 заряжает n емкостей C_k^* . Если же переключается только один транзистор, то емкость остальных («бездействующих») составляет $(n - 1) C_k$. Даже при $n = 5 \div 10$ это значение обычно в несколько раз меньше C_k^* и может не учитываться. Как видим, схема ТЛЭС в принципе обладает исключительно высоким быстродействием.

Однако, чтобы реализовать эту возможность, нужно использовать достаточно малые значения R_k (обычно менее 1 кОм). Одновременно приходится использовать достаточно большие значения тока, с тем чтобы обеспечить приемлемые значения так называемого *логического перепада*

$$\Delta U_d = |U_{„1“\text{вых}}| - |U_{„0“\text{вых}}| = I_0 R_k. \quad (15-69)$$

Увеличение тока сопровождается повышением напряжения E_0 и ростом потребляемой мощности¹. В итоге логические перепады все же получаются малыми, обычно менее 1 В, что приводит к низкой помехоустойчивости.

Основным недостатком простейшей схемы ТЛЭС является принципиальное различие однотипных логических уровней (единиц или нулей) на входе и на выходе: выходные уровни больше (по модулю) входных [см. (15-67)]. Такое различие исключает возможность работы логических элементов «друг на друга» в цепочке, что необходимо для большинства цифровых устройств.

Указанный недостаток устранен в более сложном варианте ТЛЭС (рис. 15-26) с помощью выходного повторителя T_3 . Не меняя величины логического перепада повторитель уменьшает оба выходных уровня на величину напряжения $U_{э6}$ в транзисторе T_3 . Напомним, что для кремниевых транзисторов значение $U_{э6}$ составляет около 0,7 В. Если такого смещения уровней недостаточно для «выравнивания» соответствующих выходных и входных потенциалов, используют в повторителе составной транзистор (пару Дарлингтона, см. § 4-10) или/и включают последовательно с эмиттером диод в прямом направлении. Применение повторителя помимо решения главной задачи обеспечивает существенное уменьшение выходного сопротивления, т. е. повышение нагрузочной способности схемы.

Замена резистора R_0 на транзистор T_4 , играющий роль динамической нагрузки (см. рис. 9-10), обеспечивает, во-первых, отсутствие источника питания E_0

¹ Уменьшать сопротивление R_0 с целью увеличить ток нежелательно, так как при этом коэффициент γ_6 может стать существенно меньше единицы и тогда возрастет постоянная времени τ_b [ср. (15-68) и (7-41б)].

и, во-вторых, повышение стабильности тока I_0 , что очень важно для сохранения неравенства (15-67) в диапазоне температур¹.

Замена диода D на эмиттерный переход транзистора T_5 принципиальна, однако она повышает логические возможности схемы: с коллектора T_5 снимается сигнал «1» при поступлении «1» на любую комбинацию входов, т. е. выполняется логическая функция ИЛИ. Повторитель T_0 выполняет ту же функцию, что и T_3 .

Итак, на примере ТЛЭС мы проиллюстрировали возможность использования ненасыщенных транзисторных ключей без нелинейной обратной связи. Такая возможность имеется и, как было показано, позволяет достигнуть рекордного быстродействия. Однако использование ключей ТЛЭС связано с повышенной потребляемой мощностью, малыми перепадами напряжения, а также с необходимостью обеспечивать достаточную стабильность рабочих потенциалов. Последнее, как правило, приводит к усложнению схемы. Поэтому ненасыщенные ключи (в схемах ТЛЭС) получили особое распространение лишь в микроэлектронике, где прежнее понятие «сложности» благодаря технологической интеграции качественно изменилось.

Современные схемы ТЛЭС в интегральном исполнении характеризуются временами переключения 1—2 нс, логическими перепадами 0,5—0,7 В, потребляемой мощностью в десятки милливатт.

15-8. МОЩНОСТЬ, РАССЕИВАЕМАЯ ТРАНЗИСТОРОМ В КЛЮЧЕВОМ РЕЖИМЕ

Режим отсечки характеризуется малыми токами при значительных напряжениях, а режим насыщения, наоборот, малыми напряжениями при больших токах. Последнее сочетание относится и к ненасыщенным ключам. Очевидно, что мощности, рассеиваемые транзистором в двух основных состояниях ключа, меньше,

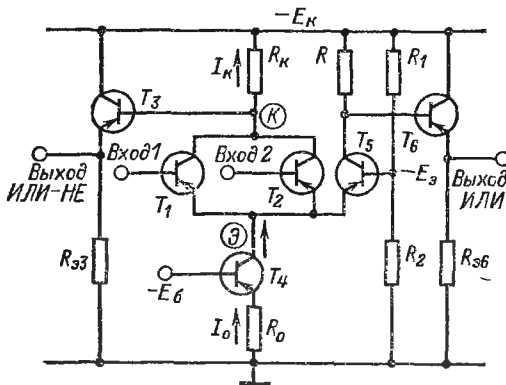


Рис. 15-26. Рабочая схема ТЛЭС с выровненными логическими уровнями на входе и выходе.

¹ В простейшей схеме изменение величин U_d и $U_{эб}$ с температурой согласно (15-64) в той или иной степени влияет на ток I_0 , тогда как в схеме на рис. 15-26 ток I_0 практически не зависит от коллекторного напряжения транзистора T_4 . Что касается непосредственного влияния температуры на ток I_{K4} , то его при необходимости можно скомпенсировать методами, рассмотренными в § 13-4.

чем в активном режиме. Отсюда следует, что *рабочий ток в ключе может значительно превышать ток, допустимый в усилительном каскаде*. Такой вывод в общем справедлив и подтверждается справочными данными. Однако при решении вопроса о допустимых токах в ключе не следует забывать о мощности, рассеиваемой при формировании фронтов, когда транзистор временно находится в активном режиме. При большой частоте переключений эта мощность играет существенную, а иногда определяющую роль. Оценим мощность во всех трех режимах ключа ОЭ.

В режиме отсечки ток коллектора равен $I_{к0}$ и примерно на порядок превышает ток эмиттера. Напряжение на коллекторном переходе составляет $E_k + E_6$ и обычно в несколько раз больше напряжения E_6 на эмиттерном переходе. Таким образом, в режиме отсечки мощность определяется коллекторной цепью и, если $E_6 < E_k$, составляет:

$$P_{отс} \approx E_k I_{к0}. \quad (15-70)$$

Например, при $I_{к0} = 5$ мкА; $E_k = 10$ В получается $P_{отс} \approx 0,05$ мВт.

В режиме насыщения напряжение U_k положительно, но меньше $U_э$. Ток I_k по-прежнему вытекает из коллектора. Это соответствует поглощению мощности на коллекторном переходе и выделению — на эмиттерном. Результирующая мощность будет равна:

$$P_{нас} \approx U_{эк} I_{к,н} + I_{к,н}^2 r. \quad (15-71)$$

Здесь принято $I_э \approx I_{к,н}$, а омические потери объединены в одном члене $I_{к,н}^2 r$, где r — эквивалентное сопротивление, которое можно оценить по формуле

$$I_{к,н}^2 r = I_6^2 r_6 + I_э^2 r_{ээ} + I_{к,н}^2 r_{кк}.$$

Практически у сплавных транзисторов $r = 2-20$ Ом, а у планарно-эпитаксиальных $r = 10-100$ Ом.

Например, если $U_{эк} = 0,1$ В; $I_{к,н} = 5$ мА и $r = 15$ Ом, то $P_{нас} \approx 1,1$ мВт. Заметим, что в данном случае мощность выделяется в основном на эмиттерном переходе¹. В ненасыщенных ключах напряжение $U_{эк}$ больше, чем в насыщенных, и зависит от э. д. с. фиксации E_ϕ (см. рис. 15-24, а). Соответственно больше будет и значение $P_{нас}$.

Во время формирования фронтов примем изменения коллекторного тока и напряжения линейными, а длительности обоих фронтов — одинаковыми. Тогда мгновенная мощность, рассеиваемая в коллекторном переходе, будет иметь вид:

$$p_k = i_k u_k = E_k I_{к,н} \frac{t}{t_\phi} \left(1 - \frac{t}{t_\phi} \right).$$

¹ У дрейфовых транзисторов, характерных большим сопротивлением коллекторного слоя (см. с. 467), основная мощность может выделяться на сопротивлении $r_{кк}$.

Интегрируя функцию p_k в пределах от 0 до t_{Φ} , деля на период переключения T и удваивая результат, получаем среднюю мощность за время обоих переключений (прямого и обратного):

$$P_{\Phi} = \frac{E_k I_{k.н}}{6} \frac{2t_{\Phi}}{T}. \quad (15-72)$$

Например, при $E_k = 10$ В; $I_{k.н} = 5$ мА; $t_{\Phi}/T = 0,05$ получается $P_{\Phi} \approx 0,85$ мВт.

Обозначая через $t_{отс}$ и $t_{нас}$ время нахождения ключа в состояниях отсечки и насыщения и используя выражения (15-70) — (15-72), запишем полную мощность, рассеиваемую в ключе:

$$P = I_{k.н}^2 \frac{t_{нас}}{T} + E_k I_{k.н} \frac{t_{отс}}{T} + U_{эк} I_{k.н} \frac{t_{нас}}{T} + \frac{1}{3} E_k I_{k.н} \frac{t_{\Phi}}{T}. \quad (15-73)$$

Эта мощность, конечно, не должна превышать допустимое значение $P_{доп}$ для данного типа транзистора. При этом в зависимости от частоты и скважности управляющего сигнала получаются различные значения допустимого тока коллектора.

Пренебрегая малосущественным слагаемым, соответствующим интервалу отсечки, получим из условия $P \leq P_{доп}$ следующее общее выражение:

$$I_{k.н} \leq \frac{E_k \theta_{\Phi} + 3U_{эк} \theta_{нас}}{6r \theta_{нас}} \left[\sqrt{\frac{36r \theta_{нас} P_{доп}}{(E_k \theta_{\Phi} + 3U_{эк} \theta_{нас})^2} + 1} - 1 \right], \quad (15-74)$$

где

$$\theta_{\Phi} = t_{\Phi}/T; \quad \theta_{нас} = t_{нас}/T.$$

Рассмотрим типичные частные случаи. Пусть, например, $\theta_{\Phi} = 0$ [практически это означает $\theta_{\Phi} \ll \theta_{нас} (3U_{эк}/E_k)$]; тогда

$$I_{k.н} \leq \frac{U_{эк}}{2r} \left[\sqrt{\frac{4r P_{доп}}{(U_{эк})^2 \theta_{нас}} + 1} - 1 \right]. \quad (15-75a)$$

Например, при $\theta_{нас} = 0,5$; $U_{эк} = 0,1$ В; $P_{доп} = 150$ мВт; $r = 15$ Ом из формулы (15-75a) получаем $I_{k.н} \leq 140$ мА. Для $\theta_{нас} = 1$ (стационарное насыщение) $I_{k.н} \leq 97$ мА.

Пусть $\theta_{\Phi} = 0,5$, т. е. импульсы имеют треугольную форму без участков насыщения и отсечки; тогда, полагая в формуле (15-74) $\theta_{нас} \rightarrow 0$ и разлагая квадратный корень в ряд, получаем:

$$I_{k.н} \leq \frac{6P_{доп}}{E_k}. \quad (15-75b)$$

Например, если $E_k = 10$ В и $P_{доп} = 150$ мВт, то $I_{k.н} \leq 90$ мА. Как видим, этот случай накладывает на ток $I_{k.н}$ наиболее сильное ограничение.

На основании изложенного можно сделать следующие выводы:

1. Допустимые токи в режиме переключения значительно (по меньшей мере в 6 раз) больше, чем в усилительном режиме [см. (15-75)].

2. В ненасыщенных ключах допустимые токи меньше, чем в насыщенных, особенно при низких частотах переключения.

3. Рассеиваемая в транзисторе мощность увеличивается, а допустимый ток уменьшается с ростом частоты переключений.

4. Мощность, рассеиваемая в режиме отсечки, не имеет существенного значения, и ее можно не учитывать даже при высокой температуре.

Глава шестнадцатая СИММЕТРИЧНЫЙ ТРИГГЕР

16-1. ВВЕДЕНИЕ

Структура симметричного транзисторного триггера (рис. 16-1) аналогична структуре лампового варианта [162—164].

В триггере на рис. 16-1 использовано независимое смещение, которое в большинстве случаев оказывается предпочтительным. Однако возможны схемы с автоматическим смещением и даже без смещения (см. § 16-3).

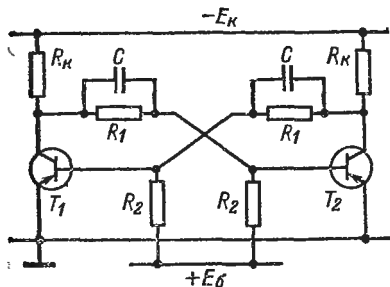


Рис. 16-1. Принципиальная схема симметричного триггера с независимым смещением.

При анализе удобно исходить из условия, что закрытое состояние одного активного элемента обуславливает открытое состояние другого и наоборот. Для определенности будем считать исходным состоянием триггера такое, в котором транзистор T_2 закрыт, а T_1 открыт. Поскольку схема симметрична, все соотношения действительны и для обратного состояния.

В дальнейших рассуждениях подразумевается насыщенный режим транзистора. Добавление элементов нелинейной обратной связи (см. § 15-7) не нарушает работы триггера, рассчитанного на насыщение. Строить же триггер так, чтобы в исходном состоянии открытый транзистор находился в активном режиме (без нелинейной обратной связи), не рекомендуется; в этом случае снижается помехоустойчивость схемы, а выходное напряжение и другие параметры будут существенно зависеть от изменений температуры.

16-2. СТАТИЧЕСКИЙ РЕЖИМ

При анализе статического режима будем сначала считать насыщенный транзистор точкой с нулевым потенциалом, а запертый транзистор — генератором тока $I_{к0}$, действующим в цепи база — коллектор.

Последняя аппроксимация основана на том, что ток эмиттера в режиме глубокой отсечки значительно меньше двух других токов

[см. (15-1)], поэтому эмиттерную цепь в таком режиме можно считать оборванной. Тогда эквивалентную схему триггера в исходном состоянии можно представить так, как показано на рис. 16-2. Воспользуемся этой схемой для вывода условий запираания и насыщения транзисторов, а также для определения выходных величин и нагрузочной способности триггера.

Условие запираания транзистора. Для того чтобы транзистор T_2 был закрыт, нужно обеспечить положительный потенциал U_{62} при наиболее неблагоприятных условиях. Обозначим запирающий потенциал через U_{60} и запишем его согласно рис. 16-2 в таком виде:

$$U_{60} = E_6 \frac{R_1}{R_1 + R_2} - I_{K0} (R_1 \parallel R_2). \quad (16-1)$$

Полагая $U_{60} > 0$, приходим к условию запираания

$$R_2 < \frac{E_6}{I_{K0}}, \quad (16-2)$$

которое с учетом допусков должно выполняться при минимальном значении E_6 и максимальном значении I_{K0} .

Обычно значение E_6 бывает задано. Если же имеется возможность выбирать значение E_6 , необходимо учитывать, что с точки зрения длительности фронтов, как будет показано ниже, нежелательно г л у б о к о е запираание транзисторов. Поэтому при расчете быстросрабатывающих триггеров следует ориентироваться на условие

$$\varphi_T \ll E_6 < (0,2 \div 0,3) E_K. \quad (16-3)$$

В случае кремниевых транзисторов током I_{K0} можно пренебречь, но зато приходится учитывать напряжение $U_{KЭ,н}$ на насыщенном транзисторе, которое нередко составляет 1 В и больше (см. § 15-2). Если на рис. 16-2 учесть потенциал $U_{K1} < 0$, то, как легко убедиться,

$$U_{60} = E_6 \frac{R_1}{R_1 + R_2} - I_{K0} (R_1 \parallel R_2) - U_{KЭ,н} \frac{R_2}{R_1 + R_2}, \quad (16-4)$$

а условие запираания будет иметь вид:

$$R_2 < \frac{E_6}{I_{K0} + U_{KЭ,н}/R_1}. \quad (16-5)$$

Как видим, для кремниевых транзисторов сопротивление R_2 нельзя выбрать независимо от сопротивления R_1 , которое оценивается на следующем этапе расчетов. Неравенство (16-3) остается в силе.

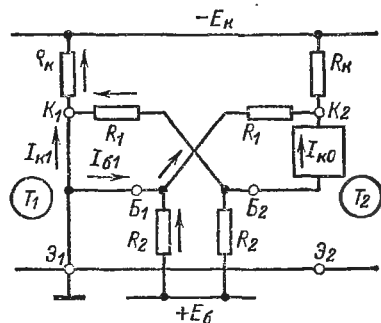


Рис. 16-2. Эквивалентная схема триггера в исходном состоянии (транзистор T_1 насыщен, транзистор T_2 заперт).

В случае дрейфовых транзисторов значение U_{60} ограничено пробоем эмиттерного перехода. Если пробой нежелателен, следует либо ограничить напряжение смещения:

$$E_6 < U_{\text{проб}} \left(1 + \frac{R_2}{R_1} \right), \quad (16-6)$$

либо ввести диодную фиксацию потенциала базы на уровне, меньшем $U_{\text{проб}}$, либо включить последовательно с базой защитный диод (см. рис. 15-22).

Условие насыщения транзистора. Базовый ток насыщенного транзистора состоит из двух компонентов: тока обратной связи, вытекающего через сопротивление R_1 , и тока, втекающего через сопротивление R_2 . Ток обратной связи легко записать, учитывая, что он обусловлен двумя источниками — генератором э. д. с. E_k и генератором тока I_{k0} . В результате получаем:

$$I_{61} = \frac{E_k - I_{k0}R_k}{R_1 + R_k} - \frac{E_6}{R_2}. \quad (16-7a)$$

Коллекторный ток транзистора T_1 тоже состоит из двух компонентов: основного нагрузочного тока, вытекающего через сопротивление R_k , и паразитного тока обратной связи, втекающего через сопротивление R_1 . Вторым компонентом, равным U_{60}/R_1 , вполне можно пренебречь, так как он на 1—2 порядка меньше первого. Таким образом,

$$I_{k1} = I_{k.n} \approx \frac{E_k}{R_k}. \quad (16-7b)$$

Подставим выражения (16-7) в неравенство (15-7). Тогда последнее после некоторых преобразований можно записать в виде *условия насыщения*

$$R_1 < R_k \left[\frac{\beta(1-\Phi)}{1 + \beta \frac{E_6/R_2}{I_{k.n}}} - 1 \right], \quad (16-8)$$

в котором величину

$$\Phi = \frac{I_{k0}}{I_{k.n}} \quad (16-9)$$

назовем *фактором теплового тока*. В случае кремниевых транзисторов можно считать $\Phi = 0$. Условие (16-8) должно выполняться при минимальных значениях R_k , β и максимальном значении Φ .

Из (16-8) следует, что сопротивление R_1 зависит в первую очередь от величины коллекторной нагрузки R_k . Последняя всегда известна в начале расчета: она выбирается исходя из желательной величины тока $I_{k.n}$ согласно (16-7b).

Если сопротивление R_2 оценивается из неравенства (16-2), то в неравенстве (16-8) величина R_2 фигурирует как заданная и выбор сопротивления R_1 не представляет затруднений. Если же

неравенство (16-2) недействительно (например, в случае кремниевых транзисторов), то сопротивления R_2 и R_1 следует определять из системы двух неравенств (16-5) и (16-8). С этой целью обозначим правые части неравенств соответственно через $R_{2\text{ макс}}$ и $R_{1\text{ макс}}$ и построим зависимости $R_{2\text{ макс}}(R_1)$ и $R_{1\text{ макс}}(R_2)$ на одном и том же графике (рис. 16-3). Эти граничные кривые определяют область допустимых значений R_1 и R_2 [160].

В случае ненасыщенного триггера с нелинейной обратной связью выбор э. д. с. E_6 и сопротивлений R_2 , R_K , R_1 производится так же, как указано выше. После этого обратная связь осуществляется по схеме, показанной на рис. 16-4 (здесь для большей ясности принято несколько иное начертание схемы, чем на рис. 16-1).

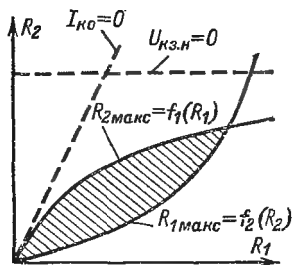


Рис. 16-3. Область допустимых значений R_1 и R_2 при их совместном выборе. Пунктиром показаны линии $R_{2\text{ макс}}$ для двух частных случаев.

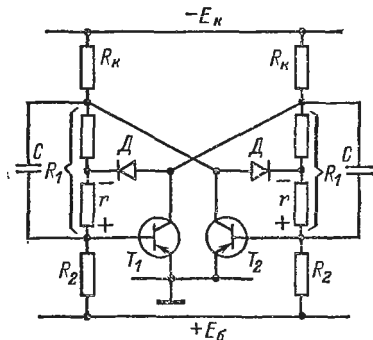


Рис. 16-4. Симметричный триггер с нелинейной обратной связью.

Легко заметить, что отличие схемы на рис. 16-4 от схемы на рис. 16-1 заключается в следующем:

1. Сопротивление R_1 разбито на две части, одна из которых (r) используется для получения э. д. с. фиксации (ср. с рис. 15-25).
2. Добавлены фиксирующие диоды D . Такое решение соответствует ключу на рис. 15-25, и сопротивление r выбирается из условия (15-47), где нужно положить $R_6 = R_2$.

Выходное напряжение и выходной ток. Выходным напряжением для схемы на рис. 16-1 является тот перепад потенциала, который получается при переходе транзистора из одного состояния в другое. Если в режиме насыщения можно принять $U_K \approx 0$, то напряжение $U_{\text{вых}}$ будет близко к потенциалу U_K запертого транзистора¹. Из рис. 16-2 легко найти U_{K2}' и, следовательно, $U_{\text{вых}}$ в таком виде:

$$U_{\text{вых}} = E_K \frac{R_1}{R_1 + R_K} (1 - \vartheta). \quad (16-10a)$$

¹ В ряде случаев приближение $U_{KЭ.н} = 0$ недействительно; тогда выходное напряжение будет соответственно меньше приводимых ниже значений (16-10).

Как видим, выходное напряжение меньше напряже-
ния питания, даже для кремниевых транзисторов, у которых $\Phi \approx$
 ≈ 0 . Если использовать (16-8) в качестве равенства и подставить
соответствующее значение R_1 в (16-10а), то выходное напряжение
можно записать в виде

$$U_{\text{вых}} = E_k \left[1 - \left(\Phi + \frac{1}{\beta} + \frac{E_6/R_2}{I_{\text{кн}}} \right) \right]. \quad (16-10б)$$

Отсюда видно, что выходное напряжение приближается к на-
пряжению E_k с увеличением тока $I_{\text{кн}}$. Обычно для ненасыщенных
схем $U_{\text{вых}} = (0,8 \div 0,9) E_k$. Влияние температуры на выход-
ное напряжение не превышает 1—2%, даже для германиевых
транзисторов.

Приведенные оценки позволяют рекомендовать выбор напряжения питания
из расчета $E_k \approx 1,1 U_{\text{вых}}$ для насыщенных схем и $E_k \approx 1,2 U_{\text{вых}}$ — для не-
насыщенных, где $U_{\text{вых}}$ — заданная величина. Разумеется э. д. с. E_k ограничена
сверху допустимым коллекторным напряжением транзистора¹. Что касается
минимального значения E_k , то оно ограничено следующими факторами. Во-пер-
вых, с уменьшением коллекторного напряжения U_k ухудшаются параметры β
и f_{α} . Во-вторых, затрудняется выполнение условий (16-3) и (16-5), обеспечивающих
крутые фронты. В-третьих, при $E_k < 2 \div 3$ В проведенный анализ делается не-
точным, так как он основан на пренебрежении межэлектродными напряжениями
насыщенного транзистора. Поэтому при использовании транзисторов в обычных
(не интегральных) схемах делают $E_k > 5 \div 10$ В. В том случае, когда необхо-
димое напряжение $U_{\text{вых}}$ меньше этих значений, его отбирают с части сопротив-
ления R_k , если это допустимо с точки зрения уровня потенциала (при непосредст-
венной связи) и значения выходного сопротивления.

Рабочий ток $I_{\text{кн}}$ не должен, разумеется, превышать значения,
допустимого в ключевом режиме (см. § 15-7). Но и слишком
малые значения тока нежелательны, а иногда недопустимы. Дей-
ствительно, уменьшение тока сопровождается уменьшением сопро-
тивлений R_k , а это затрудняет выполнение условия (16-33) и сни-
жает максимальную рабочую частоту. Кроме того, при малом токе
насыщения, как видно из (16-8), уменьшается отношение R_1/R_k ,
что согласно (16-10) приводит к уменьшению выходного напряжения.
Поэтому типичные рабочие токи триггеров в случае маломощных
высокочастотных транзисторов обычно лежат в пределах 2—5 мА.

Необходимо отметить, что благодаря наличию сопротивления
обратной связи R_1 через сопротивление R_k при запертом состоянии
транзистора протекает остаточный ток

$$I_{\text{мин}} \approx \frac{E_k}{R_k + R_1} + I_{\text{к0}}, \quad (16-11)$$

который значительно превышает значение теплового тока и может
составлять до 1 мА и выше.

¹ В связи с динамическим смещением (см. с. 495) напряжение между коллек-
тором и базой запертого транзистора в течение некоторого времени превышает
значение E_k . Поэтому целесообразно выбирать э. д. с. E_k из условия $E_k \leq$
 $\leq 0,5 U_{\text{к. доп}}$.

Статическая нагрузка. Под статической нагрузкой будем понимать активное сопротивление, присоединенное к выходу триггера гальванически (или через очень большую емкость). На рис. 16-5 показаны два основных способа подключения нагрузки, которые назовем схемами *заземленной* и *незаземленной* нагрузки.

Схема *незаземленной* нагрузки более проста для анализа, так как в ней сопротивление R_H просто шунтирует R_K и при расчетах нужно заменить R_K сопротивлением $R_K \parallel R_H < R_K$. Следовательно, при прочих равных условиях уменьшение R_H должно сопровождаться уменьшением сопротивления обратной связи в соответствии с (16-8). Величину R_1 нужно рассчитывать при минимальном значении R_H .

Главной спецификой нагруженного триггера является то, что *выходное напряжение и степень насыщения транзисторов меняются с изменением нагрузки*. При холостом ходе, когда остается только сопротивление R_K , выходное напряжение минимально, а степень насыщения максимальна; обозначим ее через $N_{х.х}$ (холостого хода). Величину $N_{х.х}$ можно найти из общего определения (15-8),

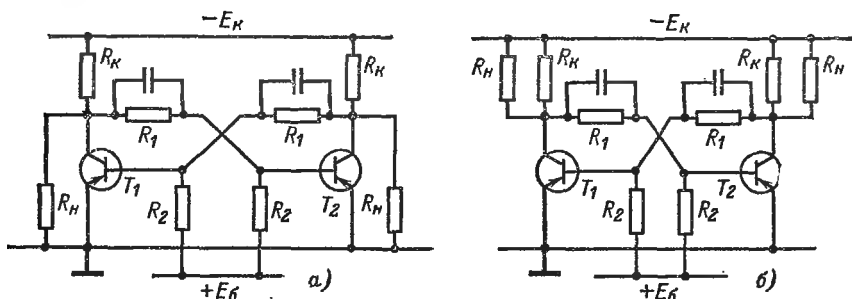


Рис. 16-5. Триггер с заземленной (а) и незаземленной (б) активными нагрузками.

подставив в него токи (16-7), причем сопротивление R_1 нужно выразить с помощью (16-8) через $R_K \parallel R_{H. \min}$. После подстановок и преобразований получим приближенные выражения:

$$N_{х.х} \approx \frac{\beta_{\min}}{1 + (\beta_{\min} - 1) \frac{R_{H. \min}}{R_{H. \min} + R_K}} - 1; \quad (16-12)$$

$$(U_{\text{вых}})_{х.х} \approx \frac{\beta_{\min} - 1}{\beta_{\min} + R_K / R_{H. \min}} E_K. \quad (16-13)$$

Например, при $R_{H. \min} = 0,1 R_K$ и $\beta_{\min} = 10$ получится $(U_{\text{вых}})_{х.х} \approx \approx 0,5 E_K$. Если такой результат неприемлем, нужно взять меньшее значение R_K . Легко заметить, что выходное напряжение будет мало зависеть от R_H , если соблюдается условие

$$R_K < \beta_{\min} R_{H. \min}. \quad (16-14)$$

Однако при этом возрастает расход мощности.

В случае *заземленной* нагрузки степень насыщения $N_{х.х}$ при холостом ходе получается несколько больше, чем (16-12), а выходное напряжение минимально при полной нагрузке и максимально при холостом ходе. А именно, при холостом ходе $U_{\text{вых}}$ выражается формулой (16-13), а при полной нагрузке — приближенной формулой

$$U_{\text{вых}} \approx \frac{\beta_{\min} - 1}{\beta_{\min} \left(1 + \frac{R_K}{R_{H. \min}} \right)} E_K. \quad (16-15)$$

Например, если $R_{н. мин} = 0,1 R_K$ и $\beta_{мин} = 10$, то $U_{вых} = 0,08 E_K$. Такой результат, разумеется, неприемлем. Следовательно, если желательно сохранять величину $U_{вых}$ при изменениях заземленной нагрузки, необходимо руководствоваться условием

$$R_K < R_{н. мин} \quad (16-16)$$

Сравнивая (16-16) и (16-14), видим, что заземленная нагрузка приводит к гораздо большему току через транзистор в режиме насыщения и соответственно к большему расходу мощности.

16-3. СХЕМНЫЕ ВАРИАНТЫ СИММЕТРИЧНОГО ТРИГГЕРА

Триггер с автоматическим смещением. Главными достоинствами этого триггера (рис. 16-6) являются наличие только одного источника питания и высокая стабильность по отношению к изменениям э. д. с. E_K . К числу его недостатков можно отнести следующие:

1. Выходное напряжение схемы на величину U_{R_3} меньше, чем при постороннем смещении (обычно на 1—2 В).

2. Минимальный потенциал коллектора (при насыщении транзистора) не равен нулю, а составляет величину U_{R_3} . Это затрудняет конструирование некоторых схем, в которых один из рабочих потенциалов должен быть близок к нулю.

3. Схема требует двух дополнительных элементов: R_3 и C_3 , причем емкость C_3 весьма большая, а в сопротивлении R_3 бесполезно расходуется мощность, составляющая до 10% и больше полной мощности триггера.

Триггер с автоматическим смещением нет необходимости анализировать так же подробно, как основную схему, так как они в общем подобны друг другу. Отметим лишь особенности расчета. Роль э. д. с. E_6 в рассматриваемом триггере играет падение напряжения U_{R_3} , равное потенциалу насыщенного транзистора U_T . Этот потенциал можно найти по формуле (15-11в), если рассматривать насыщенный транзистор как узловую точку:

$$U_T = \frac{E_K (g_K + g_{K-1})}{g_K + g_{K-1} + g_2 + g_{1-2} + g_3} \quad (16-17)$$

Здесь проводимости с двойными индексами соответствуют сопротивлениям, соединенным последовательно, например: $g_{1-2} = = (R_1 + R_2)^{-1}$.

Значением U_T , так же как и значением E_6 , задаются, исходя из соображений, изложенных в § 16-2. В дальнейших расчетах

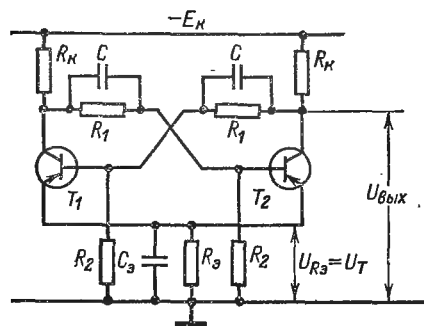


Рис. 16-6. Триггер с автоматическим смещением.

вместо э. д. с. E_6 должен фигурировать потенциал U_T . Электродвижущая сила $E'_k = E_k - U_T$, играющая в расчетах триггера с автоматическим смещением роль E_k , выбирается, как и в основной схеме, исходя из значения выходного напряжения (см. § 16-2). В конце расчета после определения всех сопротивлений находят значение R_3 из выражения (16-17).

Повышенная стабильность данной схемы объясняется тем, что напряжение смещения U_T пропорционально E_k , тогда как в основной схеме э. д. с. E_6 и E_k независимы и значительное изменение одной из них может нарушить соотношение (16-8), из которого определялось сопротивление обратной связи. Работоспособность триггера с автоматическим смещением не зависит от напряжения питания вплоть до значения $E_k = 1$ В и меньше. При этом, однако, выходное напряжение меняется пропорционально E_k .

Триггер без смещения. Способность работать без смещения (как постороннего, так и автоматического) является специфической особенностью транзисторных триггеров. Однако эти схемы на практике оказываются менее надежными. Мы дадим о них лишь самое общее представление.

На рис. 16-7 показан триггер, похожий на обычный (см. рис. 16-1), но без источника E_6 . Работоспособность такого триггера подтверждается следующими соображениями. Положим, что транзистор T_1 насыщен, и покажем, что при этом транзистор T_2 может быть «почти заперт» (т. е. хотя и будет находиться в активном режиме, но будет работать с малым коллекторным током).

В самом деле, считая транзистор T_1 точкой с нулевым потенциалом, видим, что ключ на транзисторе T_2 работает с эмиттерным переходом, зашунтированным сопротивлением $R = R_2 \parallel R_1$. Для двух предельных значений этого сопротивления легко найти соответствующие коллекторные токи. А именно, при $R = 0$, когда $U_3 = 0$, получим согласно (15-26) $I_{k2} = (1 + \beta_I) I_{k0}$; при $R = \infty$, когда $I_6 = 0$, получим согласно (4-70) $I_{k2} = (1 + \beta_N) I_{k0}^* = I_{k0}^*$. При любом конечном значении R ток I_{k2} лежит в указанных пределах. Практически можно считать $I_{k2} \approx I_{k0}^*$.

Таким образом, в данной схеме ток I_{k0}^* играет роль тока I_{k0} в схеме на рис. 16-1. Поскольку $I_{k0}^* \gg I_{k0}$, ясно, что фактор Φ увеличивается во много раз при том же токе насыщения. Тогда сопротивление R_1 согласно (16-8) сильно уменьшается, а вместе с ним согласно (16-10) уменьшается выходное напряжение.

Из сказанного ясно, что температурный диапазон данной схемы существенно ограничен, в первую очередь для германиевых транзисторов.

Так как в данной схеме $E_6 = 0$, то выражение (16-8) упрощается и принимает вид:

$$R_1 < [(1 - \Phi) \beta - 1] R_k.$$

Сопротивление R_2 желательно делать малым: при этом режим эмиттерного перехода в «запертом» транзисторе будет ближе к короткому замыканию и ток I_{k2} уменьшится по сравнению с I_{k0}^* . Однако слишком малые значения R_2 (меньше 1—2 кОм) затрудняют насыщение открытого транзистора, так как в это сопротивление отводится часть тока обратной связи.

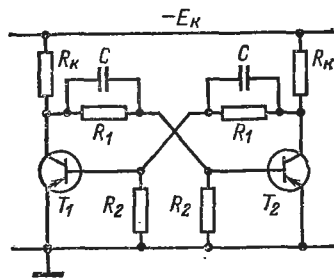


Рис. 16-7. Триггер без источника смещения.

Естественным вариантом рассмотренной схемы является триггер без сопротивлений смещения ($R_2 = \infty$). Свойства такого триггера близки к свойствам предыдущего (рис. 16-7), но триггер без сопротивлений смещения еще более критичен к изменениям температуры и коэффициента β .

Триггер с непосредственными связями. Этот вариант схемы (рис. 16-8) является наиболее оригинальным и специфичным. Он элементарно прост по структуре и весьма экономичен как в отношении числа деталей, так и в отношении потребляемой мощности. Сама возможность такой схемы вытекает из того, что уменьшение до нуля сопротивления обратной связи R_1 согласно (16-8) не нарушает условия насыщения и, следовательно, работоспособности схемы. Однако механизм работы такого триггера довольно сложен, так как он в значительной степени определяется формой вольт-амперных характеристик транзисторов в области малых токов.

Рассмотрим работу схемы. Отметим, что в данном триггере насыщенный транзистор T_1 нельзя считать эквипотенциальной точкой, так как его работа основана на разнице в межэлектродных напряжениях. Выходное напряжение, как и в других триггерах, равно разности коллекторных потенциалов транзистора в «открытом» и «закрытом» состояниях:

$$U_{\text{вых}} = U_{K2} - U_{K1}.$$

В данном случае можно подставить $U_{K2} = U_{B1}$ или $U_{K1} = U_{B2}$; тогда получается соотношение

$$U_{B1} - U_{K1} = U_{K2} - U_{B2}$$

Следовательно, $-U_{K1} = U_{K2} - U_{B2}$, т. е. напряжения U_{K1} и U_{K2} одинаковы, но имеют разные знаки (при одинаковых знаках U_{K1} схема находилась бы в совершенно симметричном состоянии и не была бы триггером). Ясно также, что схема харак-

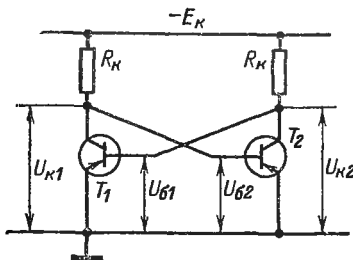


Рис. 16-8. Триггер с непосредственными связями.

терна малым выходным напряжением, так как в насыщенных транзисторах значение U_{K1} не превышает долей вольта. Таким образом, при опрокидывании данного триггера происходит в основном перераспределение токов, а напряжения меняются мало.

Поскольку транзистор T_1 насыщен, то $|U_{K1}| < |U_{B1}|$; тогда у «запертого» транзистора T_2 получается $|U_{B2}| < |U_{K2}|$. Следовательно, транзистор T_2 находится в активном режиме и слово «запертый» мы ставим в кавычки. Однако ток I_{K2} не превышает $3I_{Э0}$, так как напряжение $U_{Э02} = |U_{K1}| < \varphi_T$ [см. (15-10в)]. Токи обоих транзисторов можно найти с помощью формул (15-9) и (15-10), полагая $U_{K1} = -U_{Э02}$, $U_{K2} = -U_{Э01}$ и считая заданными и одинаковыми токи $I_{R_K} \approx E_K/R_K$. При таком расчете следует, однако, иметь в виду, что коэффициенты β при малых коллекторных напряжениях сильно отличаются от номинальных значений.

Одним из главных недостатков триггера с непосредственными связями является узкий диапазон рабочих температур (50°C и меньше). Это обусловлено примерно теми же причинами, что и у триггера без смещения.

16-4. ПЕРЕХОДНЫЕ ПРОЦЕССЫ В РЕЖИМЕ РАЗДЕЛЬНЫХ ВХОДОВ

Триггер может работать в двух существенно различных режимах: режиме раздельных входов и режиме общего — счетного входа (пересчетный режим). В первом режиме перебросы триггера достигаются либо подачей спусковых импульсов одинаковой полярности поочередно на каждый из транзисторов, либо подачей импульсов чередующейся полярности на один и тот же

транзистор. Во втором режиме импульсные токи одной полярности подаются одновременно на оба транзистора. Переходные процессы в каждом из этих режимов имеют значительную специфику. Поэтому мы рассмотрим их отдельно.

Формирование фронтов импульсов является сложным процессом. Несмотря на то что триггер является регенеративной схемой с сильной положительной обратной связью, процесс переброса определяется отнюдь не только регенерацией, т. е. не только взаимодействием обоих транзисторов. В общем случае переходный процесс состоит из трех стадий: стадии подготовки, стадии регенерации и стадии восстановления; причем стадия регенерации может занимать лишь небольшую долю всего времени переброса [147]. На первой и третьей стадиях по крайней мере один из транзисторов заперт, что существенно облегчает анализ.

Скорость опрокидывания триггера и характер переходного процесса существенно зависят от параметров спускового сигнала: его величины, формы, длительности, сопротивления источника сигнала. В данном параграфе будем считать источник входных импульсов генератором тока, а форму импульсов — прямоугольной.

Последующий анализ непосредственно относится к бездрейфовым транзисторам с не очень большой граничной частотой f_a (до 5 МГц). У таких транзисторов коллекторная емкость относительно мала, т. е. можно считать $C_k R_k \ll \tau_a$ и вместо эквивалентных постоянных времени τ_{oe} или τ_{aoe} [см. (15-32)] пользоваться величинами τ или τ_a . В случае высокочастотных дрейфовых транзисторов нужно учитывать влияние коллекторной емкости и емкости нагрузки. Это будет сделано в конце главы.

Общее описание. Пусть триггер находится в исходном состоянии (транзистор T_1 насыщен, T_2 заперт) и на вход транзистора T_1 скачком подан запирающий сигнал $I_{вх}$ (рис. 16-9). Рассмотрим переходные процессы в схеме, пользуясь временными диаграммами на рис. 16-10, на которых для простоты не учитываются токи в запертом состоянии транзисторов.

Стадия подготовки начинается в момент подачи запирающего импульса и кончается в момент отпираания транзистора T_2 .

Эта стадия состоит из двух этапов: рассасывания избыточных носителей в базе T_1 (интервал t_p) и предварительного формирования

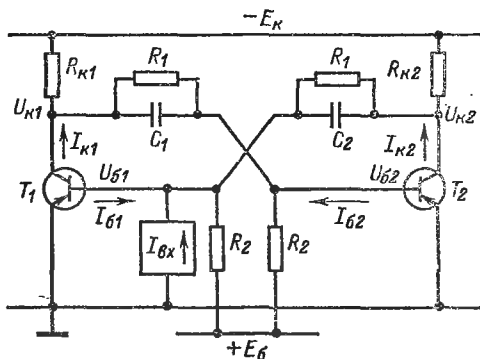


Рис. 16-9. Запуск триггера с отдельными входами.

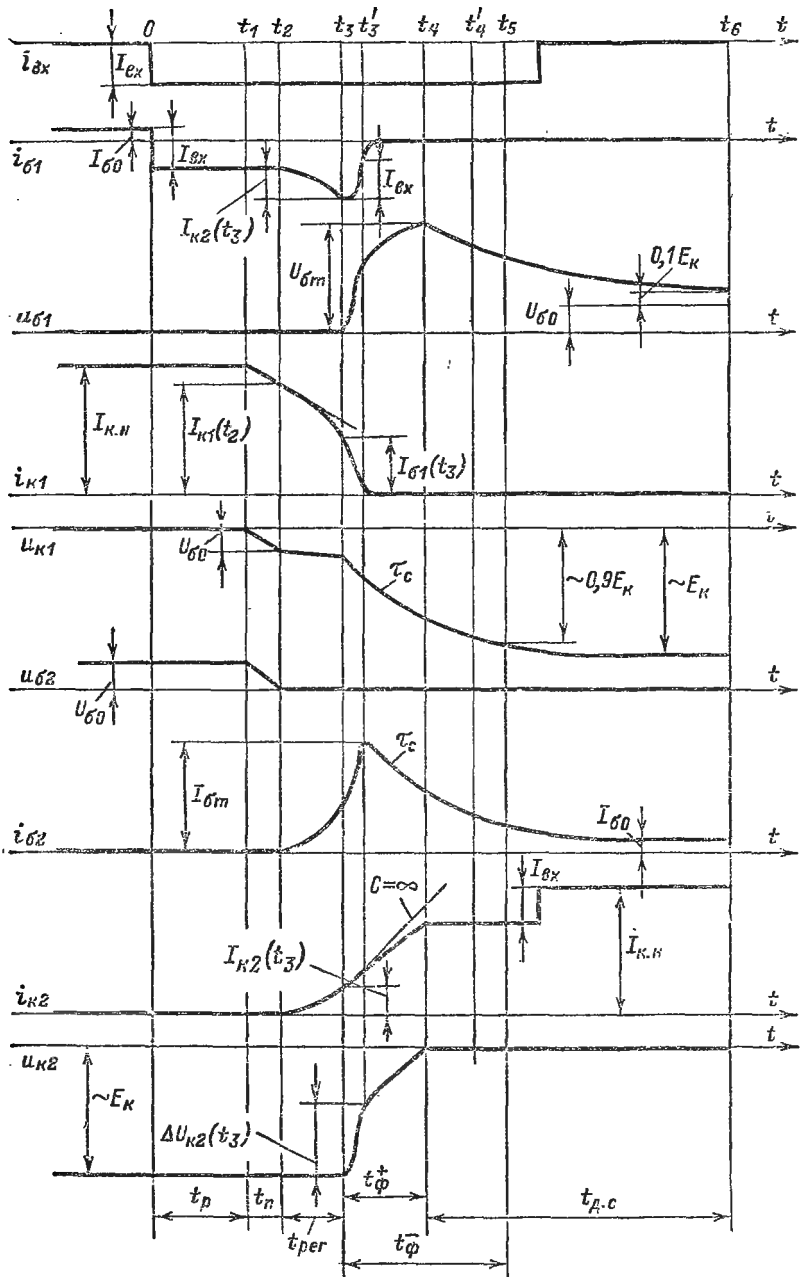


Рис. 16-10. Переходные процессы в триггере с отдельными входами.

отрицательного фронта на коллекторе T_1 (интервал t_n). На этапе t_p внешних изменений в схеме не происходит. На этапе t_n транзистор T_1 работает уже в активном режиме и ток $I_{к1}$ уменьшается. Соответственно растет напряжение $U_{к1}$. Отрицательные приращения — $\Delta U_{к1}$ передаются через конденсатор C_1 на базу транзистора T_2 и, когда приращение $\Delta U_{б2}$ полностью компенсирует начальное положительное смещение $U_{б0}$, транзистор T_2 отпирается.

Стадия регенерации начинается отпиранием транзистора T_2 , после чего в течение некоторого времени оказываются открытыми оба транзистора. Поскольку в открытом состоянии входные сопротивления малы, можно считать, что на стадии регенерации приращения коллекторных токов $I_{к1}$ и $I_{к2}$ почти полностью идут через конденсаторы C в базы смежных транзисторов. При этом, как легко убедиться, базовые токи возрастают, т. е. «форсируются» отпирание транзистора T_2 и запираание транзистора T_1 . В результате фронт тока $I_{к1}$ получается более крутым, чем при неизменном запирающем токе (последний случай показан тонкой линией).

Конец стадии регенерации определяется условием $I_{к1}(t_3) = I_{б1}(t_3)$, при котором ток $I_{э1}$ делается равным нулю. Практически в тот же момент t_3 напряжение на эмиттерном переходе становится отрицательным и транзистор T_1 оказывается в режиме динамической отсечки.

Этап динамической отсечки, строго говоря, не относится к стадии регенерации, потому что транзистор T_1 уже и е у п р а в л я е м, а значит, обратная связь отсутствует. Однако поскольку токи $I_{к1}$ и $I_{б2}$ продолжают меняться в прежнем направлении (хотя и с другой скоростью), можно условно объединить этот этап со стадией регенерации, считая, что последняя кончается в момент t'_3 , когда ток $I_{к1}$ падает до нуля, а ток $I_{б2}$ достигает максимального значения. Как известно (см. с. 490), токи транзистора в режиме динамической отсечки независимо от внешних условий падают до нуля с постоянной времени, близкой к $0,25 \tau_{\alpha}$.

Несмотря на кратковременность данного этапа, он определяет начальные условия для следующей стадии переходного процесса. На этом этапе, во-первых, заканчивается формирование импульса тока $I_{б2}$ [его полная амплитуда равна $I_{к1}(t'_3)$], а во-вторых, в результате спада тока $I_{б1}$ формируется начальный участок положительного фронта напряжения $U_{к2}$. Действительно, по мере отсечки базы ток $I_{б1}(t_3)$, равный сумме $I_{вх}$ и $I_{к2}(t_{пер})$, «вынужден» переходить в цепь $R_{к2}$ и дает на этом сопротивлении соответствующее приращение напряжения $\Delta U_{к2}(t_3)$.

Выброс базового тока $I_{бm}$ (по сравнению со статической величиной $I_{б0}$) является одним из важнейших итогов регенерации и динамической отсечки. Этот выброс, обусловленный передачей приращения $\Delta I_{к1}$ через емкость C_1 , способствует более быстрому отпиранию транзистора T_2 на следующей стадии, как и в простом ключе с форсирующей емкостью (см. § 15-4). Отсюда название емкости — *ускоряющая*. Легко заметить, что выброс $I_{бm}$ будет тем ближе к максимальному значению $I_{к.п.}$, чем меньше запирающее напряжение $U_{б0}$. Желательность слабого запираания подчеркивалась в § 16-2 и была выражена формулами (16-3) и (16-5).

Стадия восстановления начинается запирающим транзистора T_1 и кончается после достижения всеми потенциалами и токами установившихся значений, определяемых статическим режимом (T_1 заперт, T_2 насыщен). Стадия восстановления состоит из трех этапов: формирования положительного фронта на коллекторе транзистора T_2 (интервал t_{Φ}^+), формирования отрицательного фронта на коллекторе транзистора T_1 (интервал t_{Φ}^-) и этапа динамического смещения (интервал $t_{д.с.}$).

На этапе t_{Φ}^+ транзистор T_2 «доотпирается» вплоть до насыщения под действием тока $I_{б2}$. Если бы ток $I_{б2}$ не менялся и оставался равным $I_{бт}$, то формирование происходило бы, как в простом ключе, под действием весьма большого запирающего сигнала (этот случай показан тонкой линией). На самом деле ток $I_{б2}$ уменьшается в процессе формирования фронта, так как этим током заряжается емкость C_1 . Соответственно замедляется нарастание фронта и насыщение наступает несколько позднее, чем при постоянном токе базы.

На этапе t_{Φ}^- потенциал $U_{к1}$ уменьшается от значения, достигнутого в конце регенерации, до установившегося уровня, близкого к $E_{к}$. Окончание этого этапа условно, так как процесс имеет экспоненциальный характер. Постоянная времени экспоненты та же, с которой уменьшается ток $I_{б2}$, потому что оба процесса обусловлены зарядом емкости C_1 (напряжение $U_{к1}$ практически равно напряжению $U_{с1}$, поскольку $U_{б2} \approx 0$).

На этапе $t_{д.с.}$ потенциал базы закрывшегося транзистора T_1 уменьшается от того большого значения $U_{бт}$, до которого он возрос вместе с потенциалом $U_{к2}$, до статического значения $U_{б0}$. Если пренебречь разрядом емкости C_2 за время t_{Φ}^+ , то $U_{бт} \approx \Delta U_{к2} \approx E_{к}$. На самом деле емкость C_2 несколько разряжается за время t_{Φ}^+ , так что $U_{бт} < E_{к}$. Временно увеличенный запирающий потенциал базы носит название *динамического смещения* (см. с. 495).

Анализ фронтов. Из рис. 16-10 видно, что фронты импульсов в триггере имеют сложную структуру, причем положительный и отрицательный фронты формируются в разных условиях. Соответственно разными и, как увидим, противоречивыми являются способы сокращения их длительности.

Время рассасывания при запирающем импульсе $I_{вх}$ можно определить по любой из известных формул. Однако удобнее всего считать запирающий импульс сильным (что имеет место в большинстве схем) и воспользоваться выражением (15-45а) в следующем виде:

$$\theta_p \approx \frac{N}{\tilde{i}_{вх}}, \quad (16-18)$$

где $\theta = t/\tau_{\alpha}$ — относительное время; $\tilde{i}_{вх} = I_{вх}/I_{к.н}$ — относительный сигнал; $\tau_{\alpha} = \tau_n/\beta$ (это, разумеется, не очень точно, так как $\tau_n \neq \tau_p$). Сильному сигналу соответствует значение $\tilde{i}_{вх} \gg N/\beta$ (практически от 0,2 и выше).

В ненасыщенном триггере время рассасывания можно считать равным нулю, так как небольшая задержка, связанная с отбором тока из цепи обратной связи, не превышает значения τ_a (см. § 15-7).

Время предварительного формирования отрицательного фронта определяется из условия $\Delta U_{62}(t_n) = U_{60}$. При сильном сигнале начальный участок фронта можно считать линейным [см. (15-30a)]. Если, кроме того, пренебречь ответвлением тока в цепь R_2 и зарядом емкости C_1 за время T_n , то несложный анализ (см. предыдущие издания данной книги) приводит к следующим выражениям:

$$\theta_n = \frac{U_{60}/E_K}{\dot{i}_{вх}}; \quad (16-19)$$

$$I_{K1}(t_2) = I_{K.H} \left(1 - \frac{U_{60}}{E_K} \right). \quad (16-20)$$

Сравнивая (16-18) и (16-19), видим, что время t_n меньше t_p , так как $U_{60}/E_K \ll 1$, а $N > 1$.

Как правило, стадия подготовки в целом играет роль лишь при слабом сигнале (на рис. 16-10 участки t_p и t_n искусственно увеличены).

Время регенерации в общем рассчитать трудно. Однако если по-прежнему считать сигнал сильным ($\dot{i}_{вх} \gg 1/\beta$) и пренебречь зарядом ускоряющих емкостей за время t_n , то можно указать довольно простой путь решения задачи. А именно, в этом случае реакция ключа на ступенчатый сигнал имеет линейный (а не экспоненциальный) характер [см. (15-30a)]:

$$\Delta i_K(t) \approx \Delta I_6 \frac{t}{\tau_a} \approx \Delta I_6 \theta.$$

Кроме того, приращения коллекторных и базовых токов попарно равны:

$$\Delta I_{62} = \Delta I_{K1}; \quad \Delta I_{61} = \Delta I_{K2}.$$

Тогда, используя (15-30б) и учитывая начальный скачок $\Delta I_{61}(0) = I_{вх}$, можно записать следующие операторные соотношения¹:

$$\Delta I_{K1}(s) = \frac{1}{s\tau_a} [I_{вх} + I_{K2}(s)];$$

$$I_{K2}(s) = \frac{1}{s\tau_a} \Delta I_{K1}(s).$$

Из этих соотношений находим изображения коллекторных токов:

$$\Delta I_{K1}(s) = I_{вх} \frac{s\tau_a}{(s\tau_a)^2 - 1}; \quad (16-21a)$$

$$I_{K2}(s) = I_{вх} \frac{1}{(s\tau_a)^2 - 1}. \quad (16-21б)$$

¹ В начале стадии регенерации $I_{K2}(0) = 0$. Поэтому приращение ΔI_{K2} заменено на полный ток I_{K2} .

Оригиналами изображений (16-21) являются:

$$\Delta i_{k1}(\theta) = I_{\text{вх}} \operatorname{sh}(\theta); \quad (16-22a)$$

$$i_{k2}(\theta) = I_{\text{вх}} [\operatorname{ch}(\theta) - 1]. \quad (16-22б)$$

Как уже отмечалось, конец интервала регенерации определяется равенством $I_{k1}(t_3) = I_{61}(t_3)$, которое (см. рис. 16-10) записывается так:

$$I_{k1}(t_2) - \Delta I_{k1}(t_{\text{рег}}) = (I_{\text{вх}} - I_{60}) + I_{k2}(t_{\text{рег}})$$

[ток I_{60} можно найти из выражения (16-7a)].

Используя в этом равенстве формулы (16-20) и (16-22), учитывая соотношение $\operatorname{ch}(\theta) + \operatorname{sh}(\theta) = e^\theta$ и пренебрегая величинами $I_{60}/I_{\text{к.н}}$ и $U_{60}/E_{\text{к}}$ по сравнению с единицей¹, получаем простое приближенное выражение:

$$\theta_{\text{рег}} \approx \ln \frac{1}{\tilde{i}_{\text{вх}}}. \quad (16-23)$$

Это выражение действительно при $\tilde{i}_{\text{вх}} < 0,8$, что вполне естественно: если $\tilde{i}_{\text{вх}} \gtrsim 1$ (т. е. $I_{\text{вх}} \gtrsim I_{\text{к.н}}$), то запирающий ток базы практически сразу удовлетворяет условию динамической отсечки и этап регенерации отсутствует.

Подставляя (16-23) в (16-22), нетрудно найти значения токов в момент t_3 . В частности

$$I_{k2}(t_3) = \frac{1}{2} I_{\text{к.н}} (1 - \tilde{i}_{\text{вх}})^2; \quad (16-24a)$$

$$I_{61}(t_3) \approx I_{\text{вх}} + I_{k2}(t_3) \approx \frac{1}{2} I_{\text{к.н}} (1 + \tilde{i}_{\text{вх}}^2). \quad (16-24б)$$

Умножая обе части (16-24б) на $R_{\text{к}}$ и учитывая, что $I_{\text{к.н}} R_{\text{к}} = E_{\text{к}}$, получаем величину «скачка» напряжения $\Delta U_{k2}(t_3)$ (см. петит на с. 519).

Время формирования положительного фронта найдем из уже упоминавшегося условия $I_{62}(t_3) = I_{k1}(t_2)$ (см. рис. 16-10). Учитывая (16-20), запишем значение тока I_{62} в начале данного этапа:

$$I_{6m} \approx I_{\text{к.н}} \left(1 - \frac{U_{60}}{E_{\text{к}}} \right). \quad (16-25)$$

В дальнейшем ток I_{62} уменьшается экспоненциально по мере заряда емкости C_1 с постоянной времени

$$\tau_c = C(R_1 \parallel R_{\text{к}}) \approx CR_{\text{к}}. \quad (16-26)$$

Полагая $I_{6m} \gg I_{60}$, что выполняется в большинстве случаев, можно считать установившееся значение тока I_{62} равным нулю.

¹ Если такого пренебрежения не делать, то в числителе (16-23) вместо единицы будет стоять $1 + I_{60}/I_{\text{к.н}} - U_{60}/E_{\text{к}}$.

Тогда изображение этого тока примет вид:

$$I_{62}(s) = I_{6m} \frac{s\tau_c}{1 + s\tau_c}.$$

Считая такой отпирающий сигнал сильным во всем интервале t_{Φ}^+ , можем, как и раньше, считать функцию $\Delta i_{k2}(t)$ линейной; следовательно,

$$\Delta I_{k2}(s) = \frac{I_{62}(s)}{s\tau_a}.$$

Оригиналом этого изображения будет:

$$\Delta i_{k2}(\theta) = m I_{6m} \left(1 - e^{-\frac{\theta}{m}} \right), \quad (16-27)$$

где $m = \tau_c/\tau_a$ — относительная постоянная времени ускоряющей емкости.

Выражение (16-27) позволяет найти время положительного фронта, если задаться в левой части величиной $\Delta I_{k2}(\theta_{\Phi}^+)$. Эта величина согласно рис. 16-10 выражается следующим образом:

$$\Delta I_{k2}(\theta_{\Phi}^+) = I_{к.н} - I_{вх} - I_{k2}(t_3).$$

Подставляя (16-24а), нетрудно записать $\Delta I_{k2}(\theta_{\Phi}^+)$ в относительных единицах:

$$\delta = \frac{\Delta I_{k2}(\theta_{\Phi}^+)}{I_{к.н}} = \frac{1}{2} (1 - \tilde{i}_{вх}^2); \quad (16-28)$$

Используя (16-25) и (16-28) в выражении (16-27), находим время положительного фронта в виде

$$\theta_{\Phi}^+ = m \ln \frac{m}{m - \delta (1 - U_{60}/E_k)^{-1}}. \quad (16-29)$$

Из этой формулы следует, что *длительность положительного фронта сокращается при увеличении ускоряющих емкостей* (т. е. коэффициента m); в пределе (при $m \rightarrow \infty$) $\theta_{\Phi}^+ = \delta / (1 - U_{60}/E_k)$. Как видим, чем меньше δ и относительное смещение U_{60}/E_k , тем меньше время фронта. Поскольку сам коэффициент δ зависит от относительного сигнала (с изменением $\tilde{i}_{вх}$ от 0 до 1 коэффициент δ уменьшается от 1/2 до 0), желательны большие значения $I_{вх}$.

Время формирования отрицательного фронта легко найти из выражения

$$u_{к1}(t) = U_{60} + (E_k - U_{60}) \left(1 - e^{-\frac{t}{\tau_c}} \right),$$

которое описывает этот экспоненциальный фронт после окончания регенерации. Считая, что процесс заканчивается на уровне $0,9 E_k$, находим:

$$\theta_{\Phi}^- = m \ln \left[10 \left(1 - \frac{U_{60}}{E_k} \right) \right]. \quad (16-30)$$

Отсюда видно, что *длительность отрицательного фронта сокращается с уменьшением ускоряющих емкостей* (т. е. коэффициента m). Это требование противоположно тому, которое было сформулировано для положительного фронта.

Время спада динамического смещения определяется из выражения

$$u_{61}(t) = U_{6m} - (U_{6m} - U_{60}) \left(1 - e^{-\frac{t}{\tau_b}}\right),$$

где $\tau_b = C(R_1 \parallel R_2)$ — постоянная времени восстановления.

Считая, что процесс заканчивается на уровне $U_{60} + 0,1 E_k$, получаем время восстановления в следующем виде:

$$t_{д.с} = \tau_b \ln \left[10 \frac{U_{6m} - U_{60}}{E_k} \right]. \quad (16-31a)$$

Поскольку обычно $U_{60} = (0,1 \div 0,2) E_k$, а $U_{6m} = (0,5 \div 0,6) E_k$, можно считать:

$$t_{д.с} \approx \tau_b. \quad (16-31б)$$

Максимальная рабочая частота. Очередной входной импульс можно подать тогда, когда закончилось формирование наиболее длинного из фронтов; в противном случае уменьшится амплитуда выходных импульсов, а затем нарушится работа триггера. Поэтому максимальная частота входных импульсов определяется наиболее длинным из фронтов. Наибольшее значение максимальной частоты получается при равенстве переднего и заднего фронтов.

Поскольку стадии подготовки и регенерации — общие для обоих фронтов, условием их одинаковой длительности будет

$$\theta_{\Phi}^+ = \theta_{\Phi}^-.$$

Подставляя сюда (16-29) и (16-30), получаем следующее оптимальное соотношение между постоянными времени:

$$m_{\text{опт}} = \left(\frac{\tau_c}{\tau_a} \right)_{\text{опт}} = \frac{\delta}{0,9 - U_{60}/E_k}. \quad (16-32)$$

Из выражения (16-32) следует, что при больших смещениях U_{60} нужно увеличивать ускоряющие емкости C .

Так как обычно U_{60}/E_k не превышает 0,2, удобно принять для расчетов

$$\tau_c = CR_k \approx 1,5\delta\tau_a, \quad (16-33)$$

где $\delta \approx 0,5$ для входных импульсов с амплитудой $I_{\text{вх}} < 0,5 I_{\text{к.н.}}$. В общем случае значения δ рассчитываются по формуле (16-28).

Минимальный интервал между входными импульсами определяется суммой интервалов, соответствующих основным этапам переходного процесса:

$$T_{\text{мин}} = t_p + t_n + t_{\text{рег}} + t_{\Phi}^-.$$

Пренебрегая накоплением избыточных носителей, которое в принципе можно предотвратить, и интервалом τ_n , который невелик при малых относительных смещениях, получаем с учетом (16-23), (16-30) и (16-32):

$$T_{\min} \approx \tau_a \left[\ln \frac{1}{\tilde{i}_{\text{вх}}} + \delta \frac{\ln [10 (1 - U_{60}/E_k)]}{0,9 - U_{60}/E_k} \right]. \quad (16-34)$$

Если принять $U_{60}/E_k = 0 \div 0,5$; $\tilde{i}_{\text{вх}} = 0,1 \div 0,5$ и $\delta = 0,5$, то минимальный период $T_{\min}$ составляет $(2 \div 4) \tau_a$. Соответственно максимальная рабочая частота лежит в пределах

$$F_{\max} = (1,5 \div 3) f_a. \quad (16-35)$$

Под граничной частотой f_a в этом выражении следует понимать усредненное значение, которое на 20—30% меньше номинальной (справочной), поскольку коллекторное напряжение (от которого зависит f_a) изменяется и в широком диапазоне и доходит до нуля. Таким образом, практически $F_{\max} = (1 \div 2) f_{a\text{ном}}$, где $f_{a\text{ном}}$ — номинальное значение граничной частоты.

Максимальная частота уменьшается с уменьшением входного сигнала и с увеличением смещения U_{60} . Влияние температуры на максимальную частоту сказывается главным образом через напряжение U_{60} , которое растет с уменьшением температуры (см. § 16-2).

Большое динамическое смещение так же отрицательно влияет на быстродействие триггера, как и большая статическая величина U_{60} . Поэтому следует при расчете схемы обеспечить достаточный спад динамического смещения к моменту прихода очередного спускового импульса. Иначе говоря, при минимальном периоде импульсов $T_{\min}$ нужно выполнить условие

$$t_{\text{д.с}} < T_{\min} \quad (16-36)$$

или, учитывая (16-31б),

$$C(R_1 \parallel R_2) < T_{\min}. \quad (16-37)$$

При наличии диодной фиксации в базовой цепи неравенства (16-36) и (16-37) теряют смысл, поскольку в этом случае выброс напряжения U_{62} отсутствует (точнее, он не превышает напряжения фиксации, близкого к U_{60}).

16-5. ПЕРЕХОДНЫЕ ПРОЦЕССЫ В РЕЖИМЕ ОБЩЕГО ВХОДА

Упрощенная схема запуска триггера с общим входом показана на рис 16-11. Входной импульс от источника $E_{\text{вх}}$ поступает одновременно на обе базы через разделительные диоды D_1 и D_2 . По окончании входного импульса эти диоды отключают источник сигнала от триггера.

В исходном состоянии, как и раньше, транзистор T_1 считается насыщенным, а транзистор T_2 — закрытым. Диоды D_1 и D_2 в исходном состоянии закрыты отрицательным смещением E_1 .

В дальнейшем для упрощения анализа будем идеализировать диоды, пренебрегая их прямыми напряжениями, т. е. полагая, что

открытое состояние диода равносильно короткому замыканию его электродов. В общих чертах работа триггера протекает следующим образом. Положительный входной импульс, превышающий E_1 , проходит на базу транзистора T_1 и через некоторое время запирает его. В результате до окончания импульса оказываются закрытыми оба транзистора. Однако это не означает симметричного состояния схемы, так как начальное различие напряжений на емкостях сглаживается лишь постепенно. Поэтому, если входной импульс достаточно короткий, то по его окончании различие напряжений U_{c1} и U_{c2} обеспечивает должную направленность переходного процесса: отпирание транзистора T_2 и запираание транзистора T_1 , а не возврат в исходное состояние.

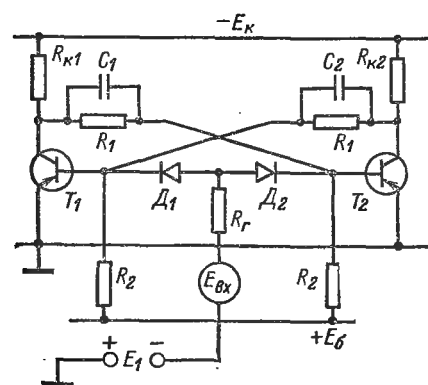


Рис. 16-11. Схема запуска триггера с общим входом (пересчетной ячейки).

и C_2 выполняют не только ускоряющую, но и еще более важную функцию. Перейдем к рассмотрению переходных процессов.

Общее описание. Начиная с поступления импульса $E_{вх}$, процессы развиваются так, как показано на рис. 16-12. В момент $t = 0$ действующее положительное напряжение $E'_{вх} = E_{вх} - E_1$ отпирает диод D_1 ; часть этого напряжения $E'_{вх}$ падает на базе 1-го транзистора. Потенциал $U_{б1}$ становится положительным и равным $E'_{вх}$; соответственно ток $I_{б1}$ резко уменьшается и становится отрицательным. Положительное приращение $\Delta U_{б1}$ через конденсатор C_2 передается на коллектор транзистора T_2 . Потенциал $U_{б2}$ либо тоже возрастает до $E'_{вх}$ (если $E'_{вх} > U_{б0}$, т.е. если стпирается диод D_2), либо не меняется и остается равным $U_{б0}$ (если $E'_{вх} < U_{б0}$, т.е. если диод D_2 не отпирается). Последний, более типичный случай показан на рис. 16-12; случай $E'_{вх} > U_{б0}$ рассмотрен в 1-м издании книги.

Весь переходный процесс, как и в режиме отдельных входов, можно разбить на несколько этапов.

Первым этапом является рассасывание избыточных носителей в транзисторе T_1 , которое, как всегда, не сопровождается внешними изменениями в схеме (интервал t_p).

Вторым этапом является предварительное формирование отрицательного фронта (интервал t_n). На этом этапе транзистор T_1 начинает запирается и ток $I_{к1}$ уменьшается. Приращения $\Delta I_{к1}$ проходят через резистор $R_{к1}$, понижая

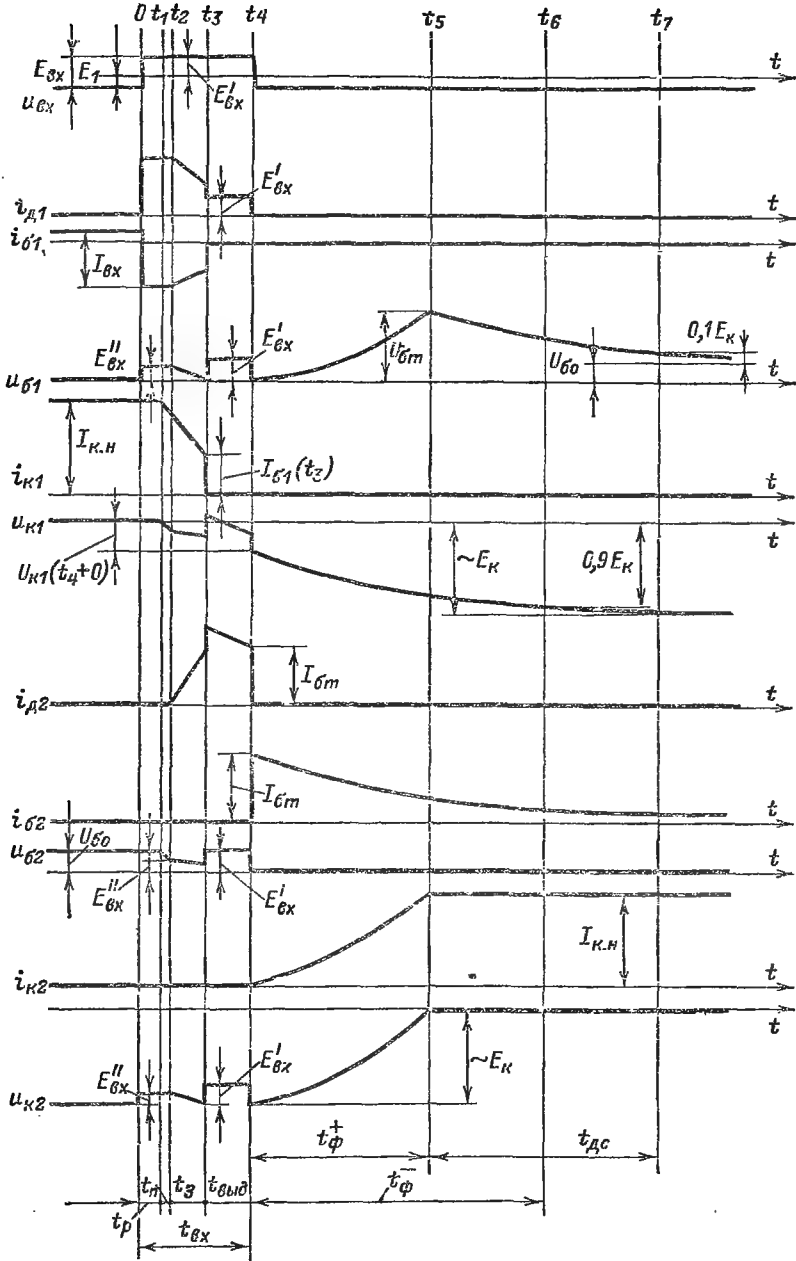


Рис. 16-12. Переходные процессы в пересчетной ячейке.

приращении $U_{к1}$. Отрицательные приращения $\Delta U_{к1}$ передаются через конденсатор C_1 на базу транзистора T_2 . Соответственно потенциал $U_{б2}$ тоже понижается и при достижении равенства $U_{б2} = E'_{вк}$ открывается диод D_2 . Этим заканчивается второй этап. Если $E'_{вк} > U_{б0}$, то диод D_2 открывается в самом начале переходного процесса и тогда данный этап отсутствует.

Третий этап можно назвать *э т а п о м з а п и р а н и я* (интервал t_3). Этап запираания формально соответствует стадии регенерации в триггере с раздельными входами. Это соответствие состоит в том, что в обоих случаях ток $I_{к1}$ падает от начального значения $I_{к1}(t_2)$ до значения $I_{к1}(t_3) = I_{б1}(t_3)$, после чего наступает этап динамической отсечки в транзисторе T_1 . Однако по существу процессы в том и другом случае *п р и н ц и п и а л ь н о* различны. Действительно, в режиме раздельных входов приращение $\Delta I_{к1}$, начиная с момента t_2 , шло в базу транзистора T_2 и вызывало нарастание тока $I_{к2}$, который, накладываясь на ток $I_{б1}$, приводил к увеличению последнего, тогда как в режиме общего входа приращение $\Delta I_{к1}$ идет через диод D_2 , частично ответвляясь в цепь базы, и протекает *н а в с т р е ч у* исходному току, т. е. приводит к *у м е н ь ш е н и ю* тока $I_{б1}$. Таким образом, *на этапе запираания в триггере с общим входом имеет место не положительная, а отрицательная обратная связь.*

Этап динамической отсечки, начинающийся в момент t_3 , характеризуется быстрым спадом токов $I_{б1}$ и $I_{к1}$ до нуля (на рис. 16-12 длительность этого этапа принята равной нулю). Спад тока $I_{б1}$ сопровождается ростом потенциалов $U_{б1}$ и $U_{б2}$ до уровня $E'_{вк}$. Скачок потенциалов $\Delta U_6 = E'_{вк} - E''_{вк}$ передается на коллекторы обоих транзисторов. Этап динамической отсечки, учитывая его малую длительность, обычно относят к этапу запираания, подобно тому, как в режиме раздельных входов его относят к стадии регенерации.

Выброс тока $I_{д2}$ в режиме общего входа является аналогом выброса тока $I_{б2}$ в режиме раздельных входов. Оба выброса имеют одинаковое происхождение и примерно одинаковую величину, близкую к $I_{к1}(t_2)$. Однако импульс $I_{б2}$ в режиме раздельных входов непосредственно обеспечивает отпирание транзистора T_2 по окончании регенерации, тогда как в режиме общего входа импульс $I_{д2}$ как бы находится в «горячем резерве» вплоть до окончания входного сигнала; только после этого он поступает в базу транзистора T_2 и обеспечивает отпирание последнего.

Четвертый этап — *э т а п в ы д е р ж к и* $t_{выд}$ — свойствен только режиму общего входа. На этом этапе, который продолжается до окончания входного импульса, оба транзистора заперты, потенциалы баз зафиксированы на уровне $E'_{вк}$, а коллекторные потенциалы меняются лишь постольку, поскольку конденсаторы C_1 и C_2 заряжаются токами $I_{д2}$ и $I_{д1}$. Ток $I_{д2}$ обычно много больше, чем $I_{д1}$, поэтому потенциал $U_{к1}$ меняется более заметно, чем $U_{к2}$.

В момент t_4 , когда кончается входной импульс, аноды диодов D_1 и D_2 снова оказываются под отрицательным потенциалом E_1 ,

а базы транзисторов — под нулевым потенциалом¹. В результате запираания диодов схема в первый момент (до появления коллекторных токов) представляет собой две независимые ветви, каждая из которых содержит эмиттерный переход соответствующего транзистора. Если пренебречь токами I_{R_2} , то базовые токи в момент t_4 определяются соотношениями

$$I_{61}(t_4) \approx [E_K - U_{c2}(t_4)] g_K; \quad (16-38a)$$

$$I_{62}(t_4) \approx [E_K - U_{c1}(t_4)] g_K. \quad (16-38b)$$

Как видим, различие токов обусловлено различием напряжений $U_c(t_4)$. Напомним, что в исходном состоянии триггера напряжения U_{c2} и U_{c1} резко различны: первое близко к E_K , а второе — к U_{60} . Если во время предшествующего интервала $t_{\text{выд}}$ эти значения изменились не сильно, то согласно (16-38) будет иметь место неравенство $I_{62}(t_4) \gg I_{61}(t_4)$. Значит, транзистор T_2 будет отпираться гораздо быстрее, а его нарастающий ток I_{K2} , поступая через емкость C_2 в базу транзистора T_1 , быстро компенсирует ток I_{61} и не дает транзистору T_1 открыться. Если конденсатор C_2 за время $t_{\text{выд}}$ зарядился до напряжения E_K или выше², то начальный ток $I_{61}(t_4)$ будет равен нулю и нет надобности его компенсировать.

Если же за время $t_{\text{выд}}$ конденсатор C_1 тоже успеет зарядиться до напряжения E_K или выше, то в момент t_4 оба транзистора окажутся запертыми, а *последующее отпирание одного из них будет в значительной мере зависеть от симметрии схемы, т. е. становится случайным*. Такие длинные входные импульсы, при которых имеет место указанная ненадежность, разумеется, недопустимы.

В момент t_4 , когда потенциалы баз уменьшаются от значения $E'_{\text{вх}}$ до нуля, на коллекторах транзисторов получаются скачки потенциалов той же величины $E'_{\text{вх}}$.

Стадия восстановления, начинающаяся в момент t_4 состоит из этапов формирования положительного и отрицательного фронтов, а также из этапа динамического смещения на базе закрывшегося транзистора. Эта стадия не имеет большой специфики по сравнению со случаем триггера с отдельными входами.

Из приведенного рассмотрения можно сделать два главных вывода:

1. Положительный фронт выходного напряжения (на рис. 16-12 это фронт потенциала U_{K2}) задержан на величину $t_{\text{вх}}$, и ему

¹ Потенциал базы не может стать существенно отрицательным, так как появляющийся при этом базовый ток протекает через соответствующий диод в обратном направлении и быстро его запирает.

² Максимальное напряжение, до которого могут зарядиться конденсаторы в интервале $t_{\text{выд}}$, составляет:

$$(E'_{\text{вх}} + E_K) \frac{R_1}{R_1 + R_K} \approx E'_{\text{вх}} + E_K.$$

Постоянная времени заряда

$$C(R_K + R_1) \parallel R_1 \approx CR_K.$$

предшествует положительный импульс, обусловленный непосредственным прохождением входного сигнала на выход через ускоряющую емкость. При большой величине сигнала указанный импульс может оказаться сравнимым с E_k , что, конечно, весьма нежелательно. Значит, величину входного сигнала $E_{вх}$ не следует делать излишне большой.

2. За время входного сигнала напряжение на конденсаторе, который в исходном состоянии связан с коллектором открытого транзистора, не должно существенно возрастать. В противном случае уменьшается отпирающий ток базы на этапе положительного фронта, а в пределе (если напряжение на конденсаторе возрастает до E_k и выше) триггер утрачивает работоспособность. Значит, длительность входного сигнала тоже не следует делать излишне большой¹. По возможности она должна быть близка к сумме $t_p + t_n + t_z$.

Анализ фронтов. Входной сигнал воздействует на схему только тогда, когда он превышает смещение E_1 . Поэтому действующим сигналом является

$$E'_{вх} = E_{вх} - E_1.$$

Импульс тока, под действием которого запирается транзистор T_1 , запишется следующим образом²:

$$I_{вх} = \Delta I_{61} \approx \frac{E'_{вх}}{R_r + r_6}. \quad (16-39)$$

Этот импульс при малом значении R_r может достигать 10 мА и больше, даже если э. д. с. $E_{вх}$ не превышает 1 В. Напряжение $E''_{вх}$, которое показано на рис. 16-12, вообще говоря, меньше действующего сигнала и составляет:

$$E''_{вх} = E'_{вх} \frac{r_6}{R_r + r_6},$$

где под r_6 понимается сопротивление участка база — эмиттер³.

¹ В работе [147] показано, что максимально допустимая длительность входного импульса может быть оценена по формуле

$$t_{вх. макс} = CR_k \ln \frac{I_{к.н}}{\Delta I_6},$$

где ΔI_6 — разность статических базовых токов транзисторов T_1 и T_2 , характеризующая асимметрию схему. Например, если $CR_k = 0,5$ мкс; $I_{к.н} = 5$ мА; $\Delta I_6 = 50$ мкА, то $t_{вх. макс} \approx 2,5$ мкс.

² Если учесть сопротивление диода, то в знаменателе добавится величина R_d , под которой следует понимать импульсное сопротивление постоянному току $(U_d/I_d)_{t=0}$.

³ В случае больших начальных токов базы такая аппроксимация не только удобна, но и практически оправдана [см. (15-15)]. В случае малых начальных токов вместо r_6 следует записать $R_{бэ}$ или $R_{вх}$; этот параметр, существенно зависящий от входного тока, неудобен для количественных расчетов.

Зная импульс $I_{\text{вх}}$ легко рассчитать время рассасывания по формуле (16-18):

$$\theta_p \approx \frac{N}{\bar{i}_{\text{вх}}}. \quad (16-40)$$

Время предварительного формирования отрицательного фронта легко найти, считая уменьшение тока $I_{\text{к1}}$ линейным. Поскольку в интервале $t_{\text{п}}$ потенциал $U_{\text{б2}}$ уменьшается от значения $U_{\text{б0}}$ до $E''_{\text{вх}}$ (а не до нуля, как в режиме раздельных входов), достаточно в формуле (16-19) заменить $U_{\text{б0}}$ на $U_{\text{б0}} - E''_{\text{вх}}$. Тогда

$$\theta_{\text{п}} = \frac{(U_{\text{б0}} - E''_{\text{вх}})/E_{\text{к}}}{\bar{i}_{\text{вх}}}. \quad (16-41)$$

Интервал $t_{\text{п}}$ пренебрежимо мал.

Время записи определяется из следующих соображений. Приращения $\Delta I_{\text{к1}}$ на данном этапе, пройдя диод D_2 , частично ответвляются в цепь базы. Обозначим ответвляющуюся часть приращения через $\gamma_6 \Delta I_{\text{к1}}$, где

$$\gamma_6 = \frac{R_{\text{г}}}{r_6 + R_{\text{г}}}$$

— коэффициент токораспределения между резисторами r_6 и $R_{\text{г}}$. Полагая $I_{\text{б1}}(0) \approx I_{\text{вх}}$, можно записать для этапа записи операторные соотношения:

$$I_{\text{б1}}(s) = I_{\text{вх}} - \gamma_6 \Delta I_{\text{к1}}(s);$$

$$\Delta I_{\text{к1}}(s) = \frac{1}{s\tau_{\alpha}} [I_{\text{вх}} - \gamma_6 \Delta I_{\text{к1}}(s)].$$

Из этих соотношений нетрудно получить оригиналы:

$$i_{\text{б1}}(\theta) = I_{\text{вх}} e^{-\gamma_6 \theta}; \quad (16-42a)$$

$$\Delta i_{\text{к1}}(\theta) = \frac{1}{\gamma_6} I_{\text{вх}} (1 - e^{-\gamma_6 \theta}). \quad (16-42б)$$

Здесь, как и раньше, $\theta = t/\tau_{\alpha}$.

Примем для простоты $I_{\text{к1}}(t_3) = I_{\text{к.н}}$. Тогда, используя условие $I_{\text{к.н}} - \Delta I_{\text{к1}}(\theta_3) = I_{\text{б1}}(\theta_3)$ (см. рис. 16-12), можно выразить время записи простой формулой:

$$\theta_3 = \frac{1}{\gamma_6} \ln \frac{1 - \gamma_6}{1 - \gamma_6/\bar{i}_{\text{вх}}}. \quad (16-43)$$

Из формулы (16-43) следует, что конечное время записи получается только при одновременном выполнении двух условий: $\bar{i}_{\text{вх}} < 1$ и $\gamma_6 < \bar{i}_{\text{вх}}$. При нарушении первого условия оказывается $\theta_3 = 0$, т. е. этап записи отсутствует и изменения всех токов происходят полностью на этапе динамической отсечки за время $0,5\tau_{\alpha}$. При нарушении второго условия оказывается $\theta_3 = \infty$; физически это означает, что из-за сильной противосвязи результирующий ток $I_{\text{б1}}$ настолько мал, что становится недействительным принятое при выводе условие сильного сигнала. Значит, «классический» этап записи реализуется лишь в случае $R_{\text{г}} < r_6$, т. е. в случае почти идеального источника э. д. с.

Время положительного фронта, как и при раздельных входах, определяется на основании формулы (16-27). В качестве начального тока I_{6m} в данном случае нужно использовать ток $I_{62}(t_4)$. Сравнивая (16-38б) с (16-25), замечаем, что отличие состоит в замене напряжения U_{60} на напряжение $U_{c1}(t_4)$. Последнее может отличаться от U_{60} в случае достаточно длинного входного импульса (см. с. 529).

Второе отличие состоит в величине приращения $\Delta I_{K2}(\theta_{\Phi}^+)$, которое для схемы с раздельными входами принципиально меньше $I_{K.H.}$, а для схемы с общим входом равно $I_{K.H.}$. В формуле (16-29) это соответствует значению $\delta = 1$.

Оценим сначала приращение $\Delta U_{c1}(t_{\text{выд}})$. Полагая для простоты ток I_{D2} постоянным и равным $I_{K.H.}$, получаем:

$$\Delta U_{c1} \approx \frac{I_{K.H.}}{C_1} t_{\text{выд}} = E_K \frac{\theta_{\text{выд}}}{m}.$$

Тогда

$$U_{c1}(t_4) = U_{c1}(0) + \Delta U_{c1} = U_{60} + E_K \frac{\theta_{\text{выд}}}{m}.$$

Подставляя это значение в (16-29) вместо U_{60} и полагая $\delta = 1$, получаем относительное время положительного фронта:

$$\theta_{\Phi}^+ = m \ln \frac{m}{m - \left(1 - \frac{U_{60}}{E_K} - \frac{\theta_{\text{выд}}}{m}\right)^{-1}}. \quad (16-44)$$

Отсюда видно, что длительность фронта увеличивается с ростом $t_{\text{выд}}$ и что время t_{Φ}^+ в схеме с общим входом больше, чем в схеме с раздельными входами, при равных значениях m и U_{60}/E_K .

Время отрицательного фронта определяется из тех же предпосылок, что и в случае раздельных входов [см. (16-30)]. Однако при раздельных входах начало этого фронта соответствует величине $U_{K1}(t_3) \approx U_{60}$, а при общем входе — величине $U_{K1}(t_4 + 0) \approx U_{c1}(t_4)$. Следовательно, в формуле (16-30) нужно тоже заменить напряжение U_{60} напряжением $U_{c1}(t_4)$, значение которого приведено выше. Тогда

$$\theta_{\Phi}^- = m \ln \left[10 \left(1 - \frac{U_{60}}{E_K} - \frac{\theta_{\text{выд}}}{m} \right) \right]. \quad (16-45)$$

Отсюда приходим к выводу, что время t_{Φ}^- в схеме с общим входом меньше, чем в схеме с раздельными входами.

Время спада динамического смещения $t_{д.с}$ можно определить по формулам (16-31).

Максимальная рабочая частота. Оценку максимальной частоты проведем, как и в § 16-4, из условия равенства значений θ_{Φ}^+ и θ_{Φ}^- . Приравнявая (16-44) и (16-45), получаем следующее оптималь-

ное соотношение между постоянными времени:

$$m_{\text{опт}} = \frac{1 + \theta_{\text{выд}}}{0,9 - U_{60}/E_k}. \quad (16-46a)$$

Как видим, правая часть зависит от длительности входного сигнала. Полагая, например, $\theta_{\text{выд}} = 1$ и $U_{60} = 0,2 E_k$, получаем:

$$(CR_k)_{\text{опт}} \approx 3\tau_a. \quad (16-46b)$$

Такое соотношение (конечно, весьма ориентировочное) примем в качестве расчетного. Сравнивая его с (16-33), видим, что при равных R_k в схеме с общим входом нужно использовать большие ускоряющие емкости.

Подставляя $m = 3$ и $\theta_{\text{выд}} = 1$ в выражение (16-45) и считая по-прежнему $U_{60} = 0,2 E_k$, находим:

$$t_{\bar{\phi}} \approx 4,5\tau_a.$$

Складывая $t_{\bar{\phi}}$ с временем выдержки $t_{\text{выд}} = n \tau_a$ и с временем запирания $t_{\bar{z}} \approx 0,7 \tau_a$ (последнее вычислено для $\gamma_6 = 0,5$ и $\bar{t}_{\text{вх}} = 0,75$), получаем минимальный период

$$T_{\text{мин}} \approx (5 + n) \tau_a. \quad (16-47)$$

Если выразить соответствующую максимальную рабочую частоту $F_{\text{макс}} = 1/T_{\text{мин}}$ через граничную частоту f_a , придем к следующему соотношению:

$$F_{\text{макс}} = \frac{2\pi}{5 + n} f_a. \quad (16-48)$$

Сравнивая (16-48) с (16-35), видим, что в схеме с общим входом быстроедействие даже при весьма коротком импульсе получается меньше, чем в схеме с отдельными входами. С увеличением длительности $t_{\text{вх}}$ это различие еще больше усиливается.

Интервал динамического смещения следует делать меньше периода $T_{\text{мин}}$, а следовательно, выбор сопротивления R_2 можно производить по уже известной формуле (16-36).

Коллекторный запуск. До сих пор считалось, что запускающие импульсы подаются на базы транзисторов. Однако можно осуществлять запуск триггеров с общим входом и через коллекторы (рис. 16-13). Будем сначала считать сопротивление $R_4 = \infty$. Тогда, как видно из рисунка, в отсутствие входных импульсов диоды D_1 и D_2 заперты, потому что на их аноды через сопротивление R_3 подан потенциал $-E_k$, а катоды имеют потенциалы $U_{k1} \approx 0$ (если транзистор T_1 открыт) и $U_{k2} = -E_k \frac{R_1}{R_1 + R_k}$ [см. (16-10a)].

Если положительный импульс $E_{\text{вх}}$ достаточно большой, а именно, если

$$E_{\text{вх}} > E_k - |U_{k2}| = E_k \frac{R_k}{R_1 + R_k},$$

то диод D_2 отпирается и входной ток проходит через него и через емкость C_2 в базу транзистора T_1 . Диод D_1 остается закрытым, так как на базу T_2 никакой сигнал при этом не поступает. Таким образом, *несмотря на общий вход, данная схема с точки зрения переходных процессов ведет себя как схема с отдельными входами* (рис. 16-9). Последнюю мы анализировали, считая запирающий входной ток заданным. В рассматриваемой схеме входной ток обусловлен действием э. д. с. $E_{вх}$ и выражается формулой (16-39), в которой роль смещения E_1 играет разность потенциалов E_k — $|U_{к2}|$, записанная выше. Зная ток $I_{вх}$, можно рассчитать время

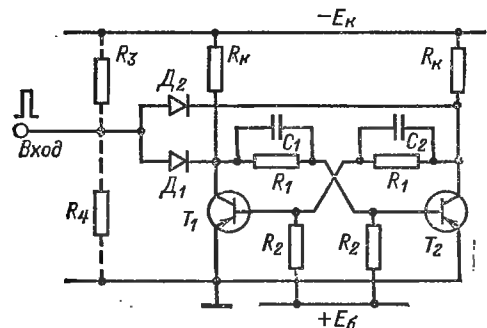


Рис. 16-13. Триггер с коллекторным запуском.

рассасывания, время предварительного формирования отрицательного фронта и время регенерации с помощью выражений (16-18), (16-19) и (16-23).

Некоторая особенность стадии регенерации состоит в том, что нарастающий ток $I_{к2}$ в принципе не полностью идет в базу транзистора T_1 : часть тока ответвляется в цепь источника сигнала через открытый диод D_2 . Однако если считать $R_k \gg r_{б1}$, то в соответствующих поправках нет необходимости.

Стадия регенерации заканчивается

запирающим транзистора T_1 , причем потенциал $U_{к1}$ остается близким к $U_{б0}$, как и в начале стадии. Запирание транзистора T_1 сопровождается уменьшением его базового тока практически до нуля. В результате входной ток триггера, т. е. ток $I_{Д2}$, который определялся сопротивлением $r_{б1}$, теперь, после отсечки базы, ограничивается гораздо большим сопротивлением R_2 и, следовательно, резко уменьшается. Ток $I_{к2}$, который до отсечки шел в базу T_1 , теперь переходит в цепь источника сигнала. Поскольку входной ток резко уменьшился, то даже небольшой доли встречного тока $I_{к2}$ достаточно для запирающего диода D_2 . Поэтому фактически диод D_2 запирается одновременно с транзистором T_1 в момент t_3 , а ток $I_{к2}$ в этот же момент полностью идет через резистор $R_{к2}$ и соответственно повышает потенциал $U_{к2}$.

Последующие процессы формирования положительного и отрицательного фронтов, а также спад динамического смещения практически не отличаются от процессов в триггере с отдельными входами (рис. 16-10). Анализ схемы с коллекторным запуском осуществляется по формулам, выведенным в § 16-4.

Можно констатировать, что коллекторный запуск триггера с общим входом имеет некоторые преимущества: отсутствует непосредственное прохождение входного импульса на коллектор отпирающего транзистора (см. кривые $u_{к2}$ на рис. 16-10 и 16-12) и отсутствует влияние длительности входного импульса на работу схемы. Последнее означает, что емкости C_1 и C_2 выполняют только ускоряющие, а не «запоминающие» функции и ограничение величины $t_{вх}$ (см. сноску ¹ на с. 530) не имеет места.

При коллекторном запуске можно использовать диоды D_1 и D_2 не только для коммутации сигнала, но и для сокращения длительности отрицательного фронта. С этой целью достаточно подключить аноды диодов к средней точке делителя R_3, R_4 и выбрать потенциал этой точки из условия

$$E_{\Phi} < E_{\kappa} \frac{R_1}{R_1 + R_{\kappa}}.$$

Тогда коллекторный потенциал запирающегося транзистора ($U_{\kappa 1}$ на рис. 16-10) будет спадать по экспоненте только до величины E_{Φ} , после чего откроется диод D_1 и «зафиксирует» потенциал $U_{\kappa 1}$ на уровне E_{Φ} . При этом длительность отрицательного фронта будет, естественно, меньше, чем тогда, когда потенциал $U_{\kappa 1}$ беспрепятственно падает до значительно большего значения $E_{\kappa} \frac{R_1}{R_1 + R_{\kappa}}$.

Особенности применения дрейфовых транзисторов. Выше при анализе триггеров подразумевалось использование бездрейфовых транзисторов. Эти транзисторы характеризуются сравнительно большими значениями постоянной времени τ_{α} и сравнительно слабым влиянием коллекторной емкости на величину эквивалентных постоянных времени (15-32): роль коллекторной емкости сводится к небольшому удлинению фронтов по сравнению с расчетными значениями; внешними «нагрузочными» емкостями обычно можно пренебречь.

Одна из важных особенностей дрейфовых транзисторов состоит в том, что у них постоянная времени τ_{α} на 2—3 порядка, а коллекторная емкость всего в 5—50 раз меньше, чем у бездрейфовых транзисторов. В результате относительная роль коллекторной емкости существенно возрастает и обычно становится определяющей. Это значит, что эквивалентные постоянные времени (15-32) близки соответственно к значениям:

$$\tau_{oe} \approx \bar{C}_{\kappa}^* R_{\kappa} \text{ и } \tau_{aoe} \approx \bar{C}_{\kappa} R_{\kappa}.$$

Если использовать значение τ_{aoe} в типовых формулах (16-33), (16-34), (16-46) и (16-47), то оптимальная величина ускоряющей емкости запишется в виде $C_{\text{опт}} = (2 \div 3) \bar{C}_{\kappa}$, а максимальная рабочая частота — в виде $F_{\text{макс}} = (0,15 \div 0,35) / \bar{C}_{\kappa} R_{\kappa}$. Например, при $\bar{C}_{\kappa} = 2$ пФ и $R_{\kappa} = 2$ кОм получаем $F_{\text{макс}} \approx 60$ МГц. В настоящее время планарные дрейфовые транзисторы обеспечивают рабочие частоты триггеров до 100—200 МГц.

Другая особенность дрейфовых транзисторов заключается в значительной роли внешних (нагрузочных и монтажных) емкостей, поскольку последние нередко оказываются сравнимыми с емкостью C_{κ} или даже превышают ее. Если при этом выполняется неравенство (15-51), то, как известно, фронты токов сокращаются, а фронты напряжений у запирающихся транзисторов описываются формулой (15-52).

ТРИГГЕР С ЭМИТТЕРНОЙ СВЯЗЬЮ

17-1. ВВЕДЕНИЕ

Триггер с эмиттерной связью (рис. 17-1) выполняет те же функции, что и его ламповый аналог — триггер с катодной связью (триггер Шмитта [124]). Будучи несимметричной схемой, триггер с эмиттерной связью используется не в качестве счетчика или запоминающего элемента подобно симметричному триггеру, а как пороговое устройство, которое реагирует на определенный уровень сигнала или определенную амплитуду импульса. В последнем

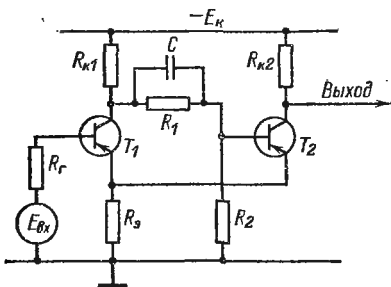


Рис. 17-1. Принципиальная схема триггера с эмиттерной связью.

случае его называют различителем или дискриминатором амплитуды. Кроме того, триггер с эмиттерной связью применяется для формирования прямоугольных импульсов определенной амплитуды из синусоидального сигнала или сигнала другой непрямоугольной формы. Таким образом, для данного триггера характерна прежде всего работа в условиях переменной величины входного сигнала.

Схемная особенность данного триггера состоит в том, что коллектор транзистора T_2 не связан с внутренними частями схемы какими-либо сопротивлениями или емкостями. Отсюда сравнительно малое влияние нагрузки на работу триггера, а также большая крутизна фронтов выходных импульсов.

Как известно, всякая спусковая схема имеет два критических уровня входного сигнала, при которых она срабатывает и отпускает, т. е. переходит из одного состояния в другое. В триггере с эмиттерной связью критические уровни имеют один и тот же знак, но разные величины. Эта разница в уровнях носит название гистерезиса.

Поскольку данный триггер чаще всего используется в качестве порогового устройства, одной из специфических проблем при его анализе и расчете является стабильность уровней срабатывания и отпускания как в зависимости от времени, так и в зависимости от температуры.

17-2. СТАТИЧЕСКИЙ РЕЖИМ

Примем условно за исходное состояние триггера такое, в котором транзистор T_1 закрыт, а T_2 открыт и насыщен. Поскольку насыщенный транзистор имеет потенциал $U_{T_2} < 0$, можно быть уверенными, что транзистор T_1 заперт, если входная э. д. с. $E_{вх} = 0$. Начиная с этого значения будем изменять э. д. с. $E_{вх}$ в сторону

отрицательных значений, а затем в обратном направлении. Тогда соответствующие изменения токов и напряжений в схеме образуют рабочий цикл триггера.

Рабочий цикл. Пренебрегая влиянием ускоряющей емкости C и считая, что источник сигнала имеет нулевое внутреннее сопротивление ($R_r = 0$), получаем кривые рабочего цикла, показанные на рис. 17-2.

Транзистор T_1 отпирается при напряжении $U_1 = U_{T2}$, после чего возрастающий ток $I_{к1}$ частично ответвляется в базу транзистора T_2 и выводит его из насыщения при напряжении U_2 . Теперь оба транзистора находятся в активном режиме, и наступает стадия регенерации. При правильном расчете схемы процесс развивается лавинообразно и происходит опрокидывание триггера во второе устойчивое состояние: транзистор T_1 насыщен, транзистор T_2 заперт. Таким образом, напряжение U_2 есть *напряжение срабатывания*. При последующем увеличении $E_{вх}$ (по модулю) транзистор T_1 все больше насыщается, так как ток $I_{с1}$ увеличивается, а $I_{к1}$ уменьшается; транзистор T_2 остается запертым. При напряжении U_3 ток $I_{к1}$ обращается в нуль, а затем делается отрицательным; при этом оба перехода транзистора T_1 работают в прямом направлении, т. е. транзистор превращается в двойной диод¹.

Если теперь уменьшать входное напряжение, то напряжения и токи в схеме изменяются по прежним кривым вплоть до точки U_2 . Обратное опрокидывание триггера происходит не в точке U_2 , а при меньшем напряжении U_3 . Получающаяся гистерезисная петля, как указывалось, характерна для любой спусковой схемы. Напряжение U_3 есть *напряжение отпускания*. Тот факт, что отпускание происходит при напряжении $U_3 < U_2$, объясняется следующим. В точке U_2 , в которой транзистор T_1 еще насыщен, потенциал $U_{б2}$ составляет лишь некоторую долю потенциала U_3 . Эта доля определяется делителем R_1, R_2 . Поэтому в точке U_2 напряжение $U_{сб2} < 0$ и транзистор T_2 находится в области глубокой отсечки. Он может открыться лишь после того, как транзистор T_1 выйдет из

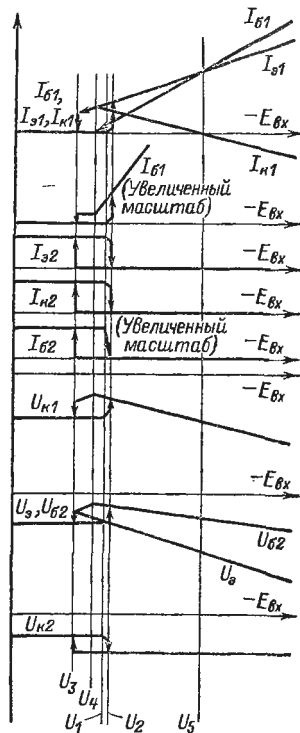


Рис. 17-2. Рабочий цикл триггера при заданной э. д. с. на входе.

¹ Такой случай, рассмотренный применительно к ключу-звезде (см. § 15-5), возможен только при условии $E_{вх} > E_{к}$, что мало реально на практике.

ниями срабатывания и отпускания, которые на рис. 17-2 обозначены через U_2 и U_3 . При конечной величине R_r линии нагрузки на рис. 17-3 должны идти с тем или иным наклоном. Тогда при малых R_r скачки тока сохраняются, но, как легко видеть из рис. 17-3, будут происходить при напряжениях, отличных от U_{II}

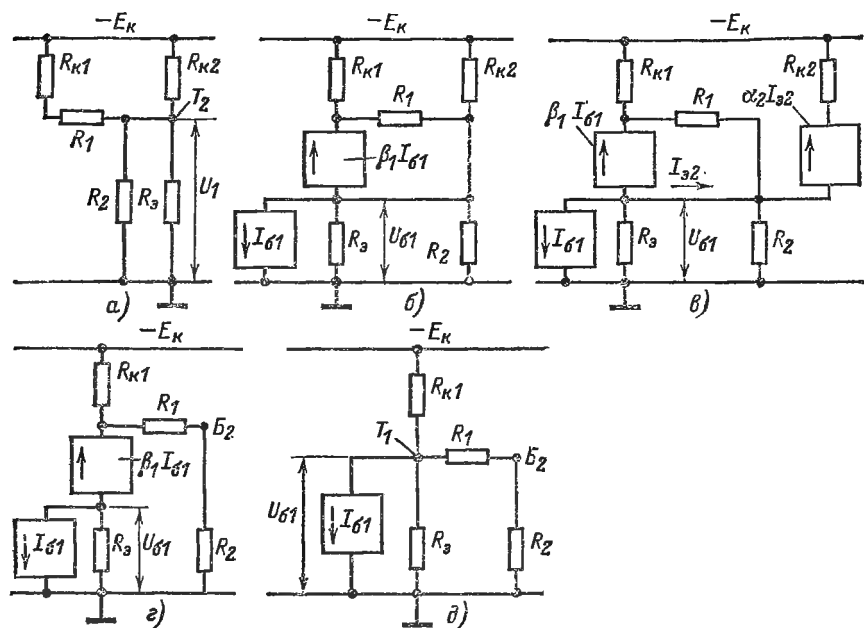


Рис. 17-4. Эквивалентные схемы триггера на основных рабочих участках.

а — начальное состояние (участок I на рис. 17-3, транзистор T_1 заперт); б — открытый транзистор T_1 , насыщенный транзистор T_2 (участок II на рис. 17-3); в — активный режим обоих транзисторов (участок III на рис. 17-3); г — открытый транзистор T_1 , запертый транзистор T_2 (участок IV на рис. 17-3); д — насыщенный транзистор T_1 (участок V на рис. 17-3).

и U_{III} . Иначе говоря, напряжения срабатывания и отпускания зависят от сопротивления R_r . Сохраняя для них прежние обозначения U_2 и U_3 , можем записать:

$$U_2 = U_{II} + I_{II} R_r; \quad (17-1a)$$

$$U_3 = U_{III} + I_{III} R_r. \quad (17-16)$$

Нетрудно заметить, что с ростом R_r рабочая точка, характеризующая схему сразу после срабатывания, перемещается с участка V на участок IV, а рабочая точка, характеризующая схему сразу после отпускания, — с участка I на участок II. При этом скачки тока уменьшаются, а напряжения U_2 и U_3 сближаются, стремясь к одинаковому максимальному значению

$$U_{2\text{ макс}} = U_{3\text{ макс}} = U_{II} + I_{II} \bar{r} = U_{III} + I_{III} \bar{r}, \quad (17-2)$$

где $\bar{r}$ — модуль отрицательного сопротивления.

Очевидно, что при условии $R_r > \bar{r}$ скачки тока будут отсутствовать и схема превратится из триггера в регенеративный усилитель. Таким образом, наличие отрицательного участка на характеристике является необходимым, но недостаточным условием триггерного режима. Ясно также, что чем больше модуль сопротивления $\bar{r}$, тем большие значения R_r допустимы при работе триггера.

Анализ входной характеристики. На рис. 17-4 показаны линейные эквивалентные схемы для каждого из участков вольт-амперной характеристики $I_{61}(U_{61})$. В этих схемах не учитываются тепловые токи и падения напряжения на открытых $p-n$ переходах. В случае необходимости поправки могут быть внесены путем добавления соответствующих генераторов тока и э. д. с. (см., например, § 17-3).

В табл. 17-1 приведены результаты анализа эквивалентных схем — вольт-амперные характеристики отдельных участков и параметры узловых точек. Анализ выполнен обычными методами теории линейных цепей с использованием условий запираия и насыщения транзисторов (подробные выводы можно найти во 2-м издании данной книги). Проводимости g в табл. 17-1 имеют те же индексы, что и сопротивления R на рис. 17-1; двойными индексами обозначены проводимости двух последовательно соединенных сопротивлений.

Главным участком входной характеристики, определяющим триггерный режим схемы, является участок III. Входное сопротивление на участке III определяем, дифференцируя (17-7):

$$\bar{r} = - \frac{dU_{61}}{dI_{61}} = \frac{\frac{1+\beta_1}{1+\beta_2} - \beta_1 \gamma_1}{g_{1-k1} + g_2 + g_3/(1+\beta_2)}. \quad (17-14)$$

Здесь $\gamma_1 = \frac{g_1}{g_1 + g_{k1}}$ — коэффициент токораспределения, который с физической точки зрения характеризует силу связи транзистора T_2 с транзистором T_1 .

Из полученного выражения видно, что сопротивление $\bar{r}$ будет отрицательным при условии

$$\beta_1 \gamma_1 > \frac{1+\beta_1}{1+\beta_2} \quad \text{или} \quad \alpha_1 \gamma_1 > (1 - \alpha_2). \quad (17-15a)$$

Для прикидочных оценок можно, полагая $\beta_1 = \beta_2$, принять

$$\beta_1 \gamma_1 > 1. \quad (17-15b)$$

Заметим, что условия (17-15) легко получить из формул (17-5) и (17-8), полагая $U_{II} > U_{III}$. Это показывает, что наличие отрицательного сопротивления на входной характеристике равносильно наличию гистерезиса на кривой рабочего цикла при $R_r = 0$.

Неравенства (17-15) являются необходимыми условиями триггерного режима. Достаточное условие триггерного режима

Основные соотношения, описывающие входную характеристику триггера с эмиттерной связью (рис. 17-1)

Участок	Вольт-амперная характеристика	Параметры узловой точки
I (рис. 17-4, а)	—	$U_I = U_{T2} = \frac{E_K}{1 + \frac{g_2 + g_3}{g_{1-K1} + g_{K2}}} \quad (17-3)$
II (рис. 17-4, б)	$U_{61} = \frac{E_K (g_{1-K1} + g_{K2}) - I_{61} [1 + \beta_1 (1 - \gamma_1)]}{g_{1-K1} + g_{K2} + g_2 + g_3} \quad (17-4)$	$U_{II} \approx \frac{E_K}{1 + \frac{g_2 + \gamma_1 g_3}{g_{1-K1} + \left(\gamma_1 - \frac{1}{\beta_2}\right) g_{K2}}} ; \quad (17-5)$ $I_{II} \approx \frac{1}{\beta_1} [E_K g_{K2} - U_{II} (g_{K2} + g_3)] \quad (17-6)$
III (рис. 17-4, в)	$U_{61} = \frac{E_K g_{1-K1} + I_{61} \left(\beta_1 \gamma_1 - \frac{1 + \beta_1}{1 + \beta_2} \right)}{g_{1-K1} + g_2 + g_3 / (1 + \beta_2)} \quad (17-7)$	$U_{III} = \frac{E_K}{1 + \frac{g_2 + \alpha_1 \gamma_1 g_3}{g_{1-K1}}} ; \quad (17-8)$ $I_{III} = -U_{III} \frac{g_3}{1 + \beta_1} ; \quad (17-9)$
IV (рис. 17-4, г)	$U_{61} = - \frac{1 + \beta_1}{g_3} I_{61} \quad (17-10)$	$U_{IV} = \frac{E_K}{1 + \frac{g_{1-2} + \alpha_1 g_3}{g_{K1}}} ; \quad (17-11)$ $I_{IV} = -U_{IV} \frac{g_3}{1 + \beta_1} \quad (17-12)$
V (рис. 17-4, д)	$U_{61} = \frac{E_K g_{K1} - I_{61}}{g_{K1} + g_3 + g_{1-2}} \quad (17-13)$	—

($R_r < \bar{r}$) с учетом (17-14) запишется в следующем виде:

$$R_r < \frac{\beta_1 \gamma_1 - \frac{1 + \beta_1}{1 + \beta_2}}{g_{1-k1} + g_2 + g_3 / (1 + \beta_2)}. \quad (17-16)$$

Отсюда следует, что если сопротивление R_r задано, то *сопротивления триггерной схемы должны быть достаточно большими*. В первую очередь это относится к сопротивлению R_{k1} , которое сильно влияет на коэффициент γ_1 .

Статический расчет. В начале расчета помимо параметров транзисторов обычно бывают заданы сопротивление R_r , напряжение питания E_k и выходное напряжение $U_{\text{вых}}$. Сопротивление R_{k2} выбирается достаточно малым с учетом переходных характеристик триггера. Величинами, подлежащими расчету, являются сопротивления R_3 , R_{k1} , R_2 , R_1 .

Напряжение $U_I = U_{T2}$ определяется из соотношения

$$U_{\text{вых}} = E_k - U_I. \quad (17-17)$$

Если заданное значение $U_{\text{вых}}$ мало, лучше ориентироваться при расчете на большее значение $(0,5 \div 0,8) E_k$, а выходное напряжение снимать с части сопротивления R_{k2} .

Сопротивление R_3 определяется из условия насыщения транзистора T_2 в исходном состоянии: $\alpha_2 I_{32} > I_{k2}$. После подстановок получается неравенство ¹

$$R_3 < \frac{\alpha_2 U_I}{U_{\text{вых}}} R_{k2}. \quad (17-18a)$$

Это неравенство не следует делать сильным во избежание значительного насыщения транзистора T_2 . Правильнее всего подставить в (17-18a) минимальные значения α_2 и R_{k2} и заменить неравенство равенством

$$R_3 = \frac{\alpha_{2\text{мин}} U_I}{U_{\text{вых}}} R_{k2\text{мин}}. \quad (17-18б)$$

Полученное значение R_3 следует считать максимальным (с учетом допусков).

Сопротивление R_{k1} определяется из условия насыщения транзистора T_1 непосредственно после срабатывания триггера. Если $R_r = 0$, то, заменяя U_{II} заведомо меньшей величиной U_I , можно записать выражение для R_{k1} в виде равенства ²

$$R_{k1} = R_{k2}. \quad (17-19a)$$

В общем случае с учетом R_r получается соотношение [165]

$$R_{k1} \approx R_{k2} \left(1 + \frac{R_r}{\beta_1 R_3} \right). \quad (17-19б)$$

Сопротивление R_2 играет ту же роль, что и в симметричном триггере с автоматическим смещением: оно обеспечивает запираание транзистора T_2 , когда тран-

¹ Подробности вывода данной и последующих формул можно найти в предыдущих изданиях книги.

² С учетом допусков целесообразно считать сопротивление R_{k1} максимальным, а R_{k2} — минимальным; это обеспечит определенную степень насыщения транзистора T_1 и позволит использовать источники сигнала с конечным сопротивлением R_r порядка сотен и даже тысяч ом.

зистор T_1 насыщен и когда в базе транзистора T_2 протекает максимальный тепловой ток $I_{к0}$. Поэтому сопротивление R_2 определяется из неравенства (16-2), если э. д. с. E_6 заменить на U_{II} . Подставив вместо U_{II} заведомо меньшую величину U_I , можно выразить R_2 следующим образом:

$$R_2 \approx \frac{U_I}{(I_{к0})_{\max}}. \quad (17-20)$$

Сопротивление R_1 обеспечивает насыщение транзистора T_2 в исходном состоянии и потому может быть определено из условия насыщения $\beta_2 I_{б2} > I_{к2}$. Подставляя соответствующие токи, нетрудно получить:

$$R_1 < R_{к2} \left(\frac{\beta_2}{1 + \beta_2 \frac{U_I/R_2}{I_{к, н2}}} - \frac{R_{к1}}{R_{к2}} \right). \quad (17-21)$$

Это выражение по форме аналогично выражению (16-8) для симметричного триггера и совпадает с ним при соблюдении условия (17-19а), т. е. при $R_1 \ll \beta_1 R_2$.

Легко убедиться, что выбор R_1 на основе формулы (17-21) гарантирует выполнение условий (17-15), необходимых для наличия участка с отрицательным сопротивлением на входной характеристике, а также условия (17-16), обеспечивающего триггерный режим при заданном значении R_1 .

В заключение поясним выбор сопротивления $R_{к2}$, которое считалось известным с самого начала расчета. В отличие от симметричного триггера, где коллекторные резисторы связаны с ускоряющими емкостями, в триггере с эмиттерной связью резистор $R_{к2}$ изолирован от емкости C . Поэтому транзистор T_2 с точки зрения переходных процессов работает в режиме, близком к режиму ключа ОЭ. Из формул (15-32) следует, что влияние коллекторной емкости на фронты импульсов становится существенным, если постоянная времени $\bar{C}_k^* R_k$ сравнима с временем жизни τ . Отсюда легко сделать вывод, что сопротивление $R_{к2}$ целесообразно выбирать из условия

$$R_{к2} < \frac{\tau}{\bar{C}_k^*} \approx \frac{\tau_\alpha}{2C_k(E_k)}, \quad (17-22)$$

где усредненная емкость $\bar{C}_k^*$ имеет величину, соответствующую отрицательному фронту [см. замечания к формулам (15-32)]. Например, если $\tau_\alpha = 0,1$ мкс и $C_k(E_k) = 5$ пФ, то $R_{к2} < 10$ кОм.

17-3. СТАБИЛЬНОСТЬ ПОРОГОВ СРАБАТЫВАНИЯ И ОТПУСКАНИЯ

Методика анализа. Анализ температурного дрейфа пороговых напряжений в такой сложной схеме, как триггер, конечно, не может быть очень строгим¹ и нуждается в упрощающих предположениях.

Будем считать источниками температурной нестабильности три параметра: тепловые токи, напряжения на открытых переходах транзисторов и коэффициенты передачи тока β . Напомним следующие особенности этих параметров. Ток $I_{к0}$ имеет заметное значение только у германиевых транзисторов при повышенной температуре. Напряжение на переходе при заданном прямом токе меняется с температурой почти линейно, причем температурная

¹ Попытка строгого анализа сделана в работе [166] на основе формул (4-4). Однако результаты такого анализа оказались малоприменимыми для практических оценок.

чувствительность отрицательна. Коэффициент β увеличивается с ростом температуры, но зависимость эта нелинейна, во всяком случае в широком диапазоне температур.

Влияние изменений температуры на работу триггера оценим с помощью входной характеристики. На рис. 17-5 сплошной линией показана характеристика при начальной (например, комнатной) температуре, а пунктирной — при некоторой другой температуре. Деформация характеристики выражается прежде всего в смещениях критических точек *II* и *III*. Для того чтобы оценить эти смещения, нет необходимости заново рассчитывать координаты точек с учетом токов $I_{к0}$ и напряжений на переходах, как сделано в работе [163]. Проще поступить следующим образом.

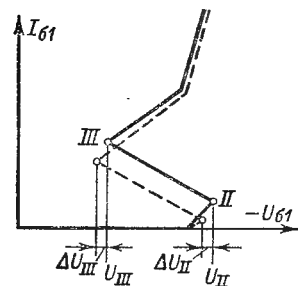


Рис. 17-5. Температурная деформация входной характеристики.

Пусть известны напряжения U_{II} , U_{III} и токи I_{II} , I_{III} для начальной температуры T_0 . Пусть теперь температура изменилась на ΔT , а существенные для стабильности параметры — соответственно на $\Delta I_{к0}$, ΔU и $\Delta \beta$. Эти приращения, если они достаточно малы, будут действовать независимо, и вызываемые ими изменения токов (ΔI_{II} , ΔI_{III}) или напряжений (ΔU_{II} , ΔU_{III}) можно сложить по принципу наложения.

Пользуясь в дальнейшем таким методом, мы не будем уточнять условий малости приращений и, таким образом, ограничимся полукачественным анализом.

Для простоты будем использовать обозначения $I_{к0}$ вместо $\Delta I_{к0}$; это не внесет большой погрешности, если принять за T_0 комнатную температуру, при которой даже у германиевых транзисторов $I_{к0} \approx 0$. Приращения напряжений на переходах будем выражать с помощью э. д. с. $e \approx \epsilon \Delta T$, где $\epsilon < 0$ и обычно составляет около 2 мВ/град (см. § 2-8). Приращения токов транзистора, вызванные изменением β , будем отражать генераторами $\Delta \alpha I_c$ или $\Delta \beta (I_6 + I_{к0})$, включенными в коллекторную цепь параллельно генераторам $I_{к0}$ или $I_{к0}^*$. Такой подход известен из гл. 6 и 13. Влияние приращений входного тока (ΔI_{II} или ΔI_{III}) будем отражать соответственно генераторами $\beta_1 \Delta I_{II}$ или $\beta_1 \Delta I_{III}$ в коллекторной цепи транзистора T_1 , как в обычном усилительном каскаде. Сопротивлениями r_k , как и раньше, пренебрежем для упрощения анализа.

Анализ температурного дрейфа. На рис. 17-6 показаны полная и упрощенные эквивалентные схемы, соответствующие участку *III* входной характеристики и отражающие влияние всех рассмотренных выше факторов [167]. Схема на рис. 17-6, б отличается от полной схемы на рис. 17-6, а объединением генераторов тока и исключением генератора э. д. с. e_1 . Последний, будучи отделен от остальной схемы генераторами тока, работает на холостом ходу и потому не может влиять на потенциал эмиттера U_e . Значит, определив

напряжение U_3 по упрощенной схеме, можно затем получить иско-
мое приращение ΔU_{61} из соотношения

$$\Delta U_{61} = \Delta U_3 + e_1. \quad (17-23)$$

Схема на рис. 17-6, а получена из предыдущей при условии $R_{K1} \approx 0$ (практически при $R_{K1} \ll R_1$, что обычно имеет место). Такая схема достаточно проста для анализа и в то же время сохраняет все основные особенности исходных схем.

Критические точки II и III входной характеристики являются границами участка III, на которых дифференциальное сопротив-

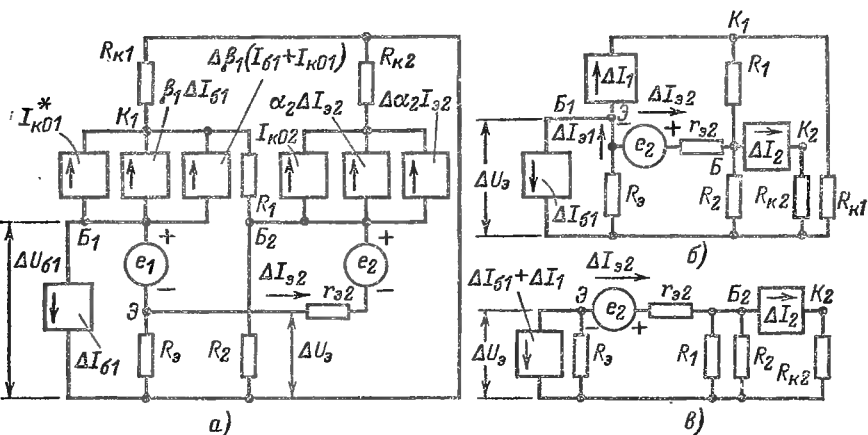


Рис. 17-6. Эквивалентные схемы триггера на участке III входной характеристики (рис. 17-3).

а — полная схема; б и в — упрощенные схемы.

ление равно нулю. Следовательно, согласно (17-14) в этих точках должно соблюдаться условие

$$\beta_2 \gamma_1 \approx 1 \quad (17-24)$$

(если $\beta_1 \approx \beta_2$).

Для соблюдения этого условия величина β_2 должна быть достаточно малой, т. е. транзистор T_2 должен быть либо достаточно сильно заперт (точка III), либо достаточно сильно насыщен (точка II), потому что в нормальном активном режиме (участок III) выполняется неравенство (17-156). Таким образом, выполнение условия (17-24) предполагает зависимость коэффициента β_2 от режима, в первую очередь от тока I_3 . Поскольку β_2 зависит еще и от температуры, полное приращение $\Delta \beta_2$ можно записать в виде

$$\Delta \beta_2 = \frac{\partial \beta_2}{\partial T} \Delta T + \frac{\partial \beta_2}{\partial I_{32}} \Delta I_{32} = \Delta \beta_T + B_2 \Delta I_{32}, \quad (17-25)$$

где $\Delta \beta_T$ — температурное приращение, а $B_2 = \partial \beta_2 / \partial I_{32}$ — чувстви-
тельность коэффициента β_2 к изменениям эмиттерного тока. Зна-

чение B_2 специально измеряется вблизи границы отсечки и насыщения. Например, для некоторых транзисторов при токах 1—10 мкА были найдены значения $B_2 = 2 \div 0,1 \text{ мкА}^{-1}$.

Из условия (17-24) следует, что вблизи критических точек полное приращение $\Delta\beta_2 = 0$, откуда получается важное соотношение:

$$\Delta I_{э2} = -\frac{\Delta\beta_T}{B_2}. \quad (17-26)$$

Это соотношение означает, что вблизи критических точек температурному приращению коэффициента β_2 должно соответствовать определенное приращение эмиттерного тока. Такое

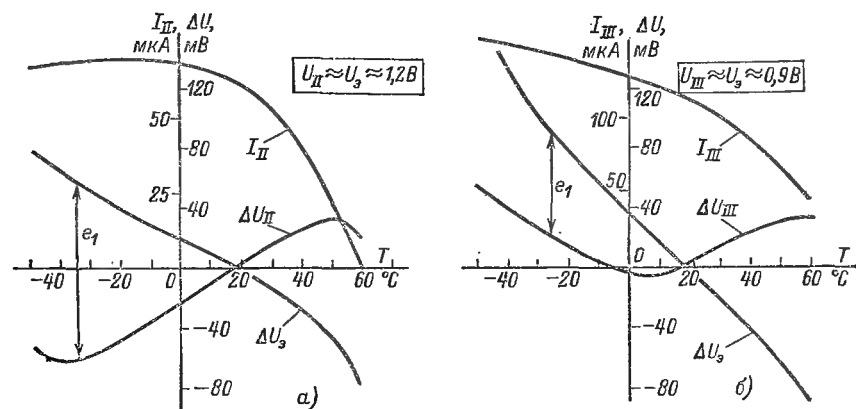


Рис. 17-7. Температурный дрейф тока и напряжения в точках срабатывания (а) и отпуска (б) в случае германиевых транзисторов.

приращение частично обеспечивается благодаря изменению внутренних факторов e , β и $I_{к0}$, зависящих от температуры. Однако в общем случае требуется менять еще и внешнюю величину — входной ток $I_{б1}$. Тогда меняются обе координаты критической точки. Если же внутренние факторы сами по себе обеспечивают приращение $\Delta I_{э2}$, требуемое соотношением (17-26), то входной ток не нужно менять: $\Delta I_{б1} = 0$. В этом случае в критической точке имеет место только дрейф напряжения.

С помощью эквивалентной схемы на рис. 17-6, в можно получить общие выражения для дрейфа критических напряжений и критических токов (см. предыдущие издания данной книги):

$$\Delta U_{б1} = -(R_1 \parallel R_2) \left[\frac{1 - \alpha_2 \gamma_{э2}}{\gamma_{э2}} \frac{\Delta\beta_2}{B_2} + I_{к02} \right] - e_2 + e_1; \quad (17-27)$$

$$\Delta I_{б1} = - \left[\frac{\Delta U_{э}}{\beta_1 R_э} + \frac{\Delta\beta_1}{\beta_1} (I_{б1} + I_{к01}) - \frac{\Delta\beta_2}{\beta_1 B_2} + I_{к01} \right], \quad (17-28)$$

где $\Delta U_{э} = \Delta U_{б1} - e_1$.

Полученные выражения вместе с экспериментальными зависимостями, показанными на рис. 17-7 и 17-8, позволяют сделать следующие качественные выводы.

1. Приращение эмиттерного потенциала всегда отрицательно, т. е. напряжение $U_{\text{э}}$ увеличивается (по модулю) с ростом температуры. Зависимость $\Delta U_{\text{э}}(T)$ нелинейна; нелинейность обусловлена в первую очередь функцией $\beta(T)$, поскольку функции $e(T)$ почти линейны, а токи $I_{\text{к0}}$ при пониженных температурах незначительны.

2. Величина и знак приращений $\Delta U_{\text{б1}}$ зависят от соотношения между модулями напряжений $\Delta U_{\text{э}}$ и e_1 , поскольку знаки последних различны. Нелинейность функции $\Delta U_{\text{э}}(T)$ делает возможными экстремумы на кривых $\Delta U_{\text{б1}}(T)$.

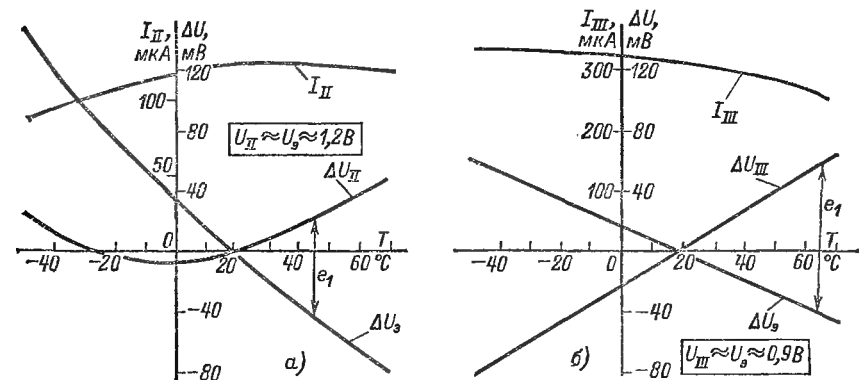


Рис. 17-8. Температурный дрейф тока и напряжения в точках срабатывания (а) и отпускания (б) в случае кремниевых транзисторов.

Для германиевых транзисторов таких экстремумов может быть два; один из них (максимум) обусловлен увеличением теплового тока при повышенной температуре.

3. Приращение $\Delta I_{\text{б1}}$ в общем случае также представляется экстремальной кривой, имеющей максимум, хотя этот максимум не всегда ярко выражен и может быть расположен вне рабочего диапазона температуры.

Резкое уменьшение тока $I_{\text{б1}}$ при повышенной температуре для германиевых транзисторов обусловлено, конечно, ростом тепловых токов и приводит к потере работоспособности схем, когда ток базы падает до нуля и напряжения U_{II} и U_{III} сближаются.

4. Кремниевые транзисторы не имеют существенных преимуществ перед германиевыми в отношении величины температурного дрейфа, поскольку главной причиной дрейфа являются не тепловые токи, а нестабильность коэффициентов передачи тока и напряжений на переходах,

Рост теплового тока у германиевых транзисторов в некоторых случаях уменьшает дрейф критических напряжений, т. е. играет положительную роль.

5. Температурный дрейф зависит от диапазона рабочих температур, типа транзисторов, их параметров и схемы; в среднем значения температурного дрейфа критических напряжений и токов составляют соответственно 0,5—1,5 мВ/°С и 0,5—1,0 мкА/°С.

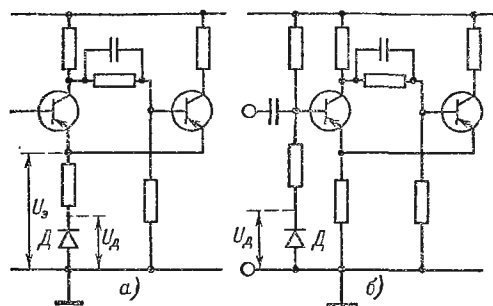


Рис. 17-9. Способы компенсации температурного дрейфа.

а — включение термистора в эмиттерную цепь;
б — включение термистора в цепь базового делителя.

Компенсация температурного дрейфа. Если бы главным источником неустойчивости были тепловые токи $I_{к0}$, то эффективным способом стабилизации триггера могло бы быть увеличение рабочих токов настолько, чтобы тепловые токи по сравнению с ними были малы даже при высокой температуре. Однако основной причиной неустойчивости являются не тепловые токи, и потому сильноточные схемы не могут быть универсальным решением, тем более что в них резко повышается потребляемая мощность.

Два основных способа, применяемых в триггерах с эмиттерной связью для частичной компенсации температурного дрейфа порогов срабатывания и отпускания, показаны на рис. 17-9. В схеме на рис. 17-9, а последовательно с резистором $R_э$ включен диод в прямом направлении. Напряжение на нем уменьшается с ростом температуры, что в значительной мере стабилизирует потенциал $U_э$, а следовательно, и критические напряжения U_{II} и U_{III} . Этот метод может иметь ряд вариантов, позволяющих осуществить более точную компенсацию. Можно, например, включать последовательно два диода или заменить $R_э$ двумя параллельно соединенными резисторами и включить диоды в одну или обе ветви.

В схеме на рис. 17-9, б использован делитель напряжения на входе триггера. Этот делитель обеспечивает запирающее транзистора T_1 в исходном состоянии. Для отпирающего этого транзистора

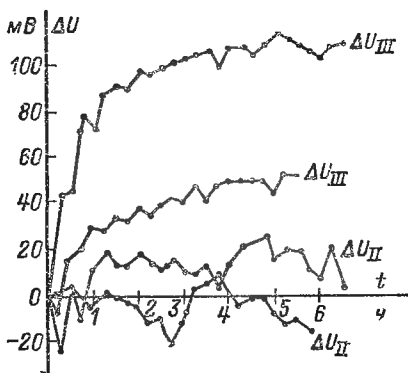


Рис. 17-10. Временной дрейф пороговых напряжений триггера.

В схеме на рис. 17-9, б использован делитель напряжения на входе триггера. Этот делитель обеспечивает запирающее транзистора T_1 в исходном состоянии. Для отпирающего этого транзистора

стора и срабатывания триггера требуется входной импульс $U_{вх} > |U_{61} - U_{11}|$. Чтобы минимальная величина спускового импульса мало зависела от температуры, нужно обеспечить одинаковые изменения потенциалов U_{61} и U_{11} . Из них изменение $\Delta U_{11}(T)$ является функцией заданной, а функцию $\Delta U_{61}(T)$ (по возможности аналогичного вида) подбирают, включая в делитель температурно-зависимые элементы — термисторы или полупроводниковые диоды в тех или иных комбинациях.

Оба указанных способа могут обеспечить уменьшение температурного дрейфа в несколько раз.

Временной дрейф. Этот вид нестабильности определяется сложными физико-химическими процессами на поверхности полупроводника, а также низкочастотной составляющей собственных шумов. Результаты измерений временного дрейфа критических напряжений показаны на рис. 17-10. Как видим, в общем случае дрейф состоит из медленной (регулярной) и более быстрой (хаотической) составляющих, причем соотношение между ними зависит от типов и даже экземпляров транзисторов. Значение регулярного дрейфа (ползучести) может достигать до 100 мВ с постоянной времени, составляющей несколько часов, а хаотический дрейф обычно не превышает 10—20 мВ, т. е. 1—2% критического напряжения.

Глава восемнадцатая

МУЛЬТИВИБРАТОРЫ

18-1. СИММЕТРИЧНЫЕ МУЛЬТИВИБРАТОРЫ

Симметричный (по структуре) мультивибратор, показанный на рис. 18-1, получается из симметричного триггера (см. рис. 16-1) путем исключения резисторов R_1 и замены смещения E_6 смещением $-E_k$. Поскольку этапы рассасывания и регенерации протекают в обеих схемах аналогично¹ и были подробно изучены применительно к триггеру, будем в дальнейшем считать изменения токов в транзисторах мгновенными. Это позволит нам сосредоточить внимание на специфических для мультивибраторов процессах.

Рабочий цикл. Работа мультивибратора в течение одного полного периода описывается временными диаграммами на рис. 18-2. Открытый транзистор, как и в триггерах, будем считать насыщенным. Очередное опрокидывание схемы происходит тогда, когда напряжение на базе ранее запертого транзистора делается равным нулю и он отпирается. При этом возрастающий коллекторный ток идет в базу другого транзистора через емкость и выводит его из насыщения, после чего наступает регенерация.

¹ Это не исключает и некоторых особенностей переходного процесса в мультивибраторах. Частично эти особенности рассмотрены в работе [168].

Пусть в момент $t = 0$ открылся транзистор T_1 , поскольку потенци-
циал $U_{б1}$, уменьшаясь, достиг нулевого значения. В результате
насыщения транзистора T_1 потенциал $U_{к1}$, который до скачка был
равен $E_k - I_{к0}R_{к1}$, падает до нуля; транзистор T_1 «стягивается
в точку». Напряжение $U_{с1}$ на конденсаторе C_1 в первый момент
не меняется и остается равным $E_k - I_{к0}R_{к1}$ (так как потенциал $U_{с2}$
до скачка был равен нулю, а потенциал $U_{к1}$ был равен $E_k - I_{к0}R_{к1}$).
Отсюда следует, что ток $I_{к1}$ в момент $t = 0$ возрастет от значе-
ния $I_{к0}$ до $E_k (g_{к1} + 2g_1)$. Слагаемое $2E_k g_1$ обусловлено тем, что
резистор R_1 после скачка находится под напряжением $U_{с1} + E_k \approx$
 $\approx 2E_k$. Напряжение на конденсаторе C_2 перед скачком было близко
к нулю и сохраняет это значение непосредственно после скачка.

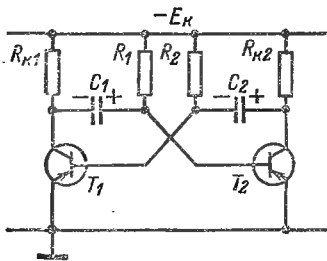


Рис. 18-1. Принципиальная
схема симметричного мультиви-
братора.

Отсюда следует, что базовый ток $I_{б1}$
(который до опрокидывания был ра-
вен $-I_{к0}$) возрастает до $E_k (g_{к2} +$
 $+ g_2)$, поскольку резисторы $R_{к2}$ и
 R_2 в первый момент соединены па-
раллельно через незаряженный кон-
денсатор C_2 . Положительный скачок
потенциала $U_{к1}$ передается через кон-
денсатор C_1 на базу транзистора T_2
и поддерживает его в запертом со-
стоянии.

Коллекторный ток $I_{к2}$ до скач-
ка был равен $E_k (g_{к2} + g_2)$, а после
опрокидывания падает до значе-
ния $I_{к0}$. Легко заметить, что ток
 $I_{к2} (-0)$ равен току $I_{б1} (+0)$, т. е. во время скачка коллектор-
ный ток запирающегося транзистора «переходит» в базу отпи-
рающегося транзистора. Потенциал $U_{к2}$ в момент $t = 0$ не изме-
няется и остается близким к нулю. Это следует из равенства
 $U_{к2} (0) = U_{б1} (+0) + U_{с2} (0)$, где оба слагаемых в правой части
очень малы. Ток $I_{б2}$, который до опрокидывания был равен $E_k g_1$,
после опрокидывания уменьшается до $-I_{к0}$. Таков в первом при-
ближении результат происшедшего переброса в схеме.

Начиная с момента $t = 0$ величины в схеме изменяются сле-
дующим образом. Потенциалы $U_{б1}$, $U_{к1}$ и токи $I_{к2}$, $I_{б2}$ остаются
практически неизменными, близкими к нулю. Потенциал $U_{с2}$,
равный напряжению на конденсаторе C_1 , уменьшается с постоян-
ной времени $C_1 R_1$, стремясь к величине $-(E_k + I_{к0} R_1)$ (послед-
няя легко получается, если мысленно убрать из схемы запертый
транзистор T_2 и заменить его со стороны базы генератором тепло-
вого тока¹). На самом деле разряд конденсатора прерывается

¹ Помимо теплового тока, который у кремниевых транзисторов практически
отсутствует, следовало бы учесть еще сопротивление, свойственное запертому
переходу и обусловленное утечками, током термогенерации (зависящим от напря-
жения) и другими причинами. Методика учета этого сопротивления будет показана
в § 20-2 применительно к блокинг-генератору.

по достижении потенциального значения U_{62} нулевого значения. Тогда отпирается транзистор T_2 и происходит обратное опрокидывание схемы. Кривая U_{62} пересекает ось времени весьма круто, так что момент переброса хорошо фиксирован. В момент обратного опрокидывания при $t = T_1$ напряжение $U_{c1}(T_1) \approx 0$. Ток I_{k1} немного уменьшается по мере разряда конденсатора C_1 , стремясь к установившемуся значению $E_K g_{k1}$. Поскольку $U_{c1}(T_1) \approx 0$, ток I_{k1} в конце первого полупериода будет равен $E_K(g_{k1} + g_1)$, т. е. уменьшится за время T_1 на величину $E_K g_1$. Ток I_{61} значительно уменьшается по мере заряда конденсатора C_2 , стремясь к установившемуся значению $E_K g_2$; последнее должно быть достаточным для того, чтобы сохранилось насыщение. Заряд конденсатора C_2 происходит с постоянной времени $C_2 R_{k2}$ и обычно заканчивается задолго до конца полупериода. Потенциал U_{k2} , равный напряжению на конденсаторе C_2 , спадает экспоненциально до установившейся величины $E_K = -I_{k0} R_{k2}$.

Второй полупериод работы схемы (T_2) не нуждается в дополнительных пояснениях (в случае $T_1 \neq T_2$ термин **полупериод**, разумеется, условен и означает соответствующую часть периода). В момент $T = T_1 + T_2$ рабочий цикл схемы начинает повторяться.

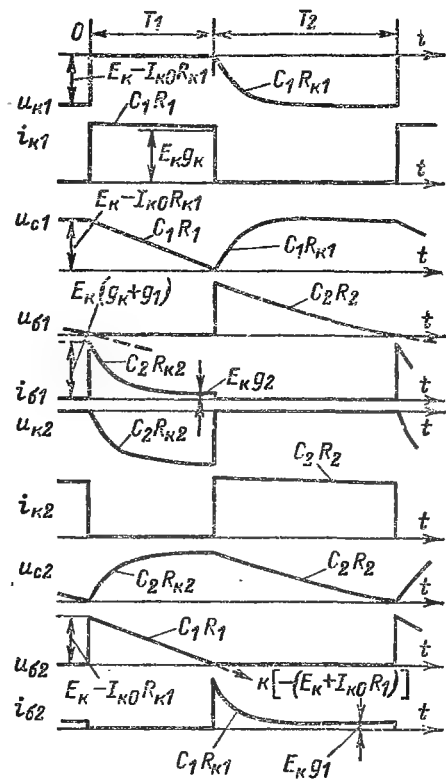


Рис. 18-2. Временные диаграммы работы мультивибратора.

Мы рассмотрели рабочий цикл, считая, что колебания уже существуют. Однако возможны случаи, когда оба транзистора мультивибратора после его включения находятся в насыщенном состоянии и колебания отсутствуют. Такой режим в генераторах синусоидальных колебаний называют режимом жесткого самовозбуждения. Причиной устойчивости симметричного начального состояния схемы являются малые коэффициенты усиления транзисторов в режиме насыщения. В результате всегда имеющиеся флуктуации тока в замкнутом контуре обратной связи могут не усиливаться, а ослабляться и будут неспособны положить начало колебательному процессу. Во избежание режима жесткого самовозбуждения следует предотвращать сильное насыщение транзисторов, чтобы сохранить доста-

точное усиление в контуре обратной связи. Нередко приходится принимать специальные меры для обеспечения асимметрии схемы при первоначальном включении источника питания [168].

Рабочая частота и ее стабильность. Выше было показано, что минимальный базовый ток насыщенного транзистора получается в конце полупериода и равен E_k/R ; коллекторный ток при этом близок к E_k/R_k . Полагая β и R_k одинаковыми у обоих транзисторов, что обычно имеет место, получаем из условия насыщения ($\beta I_6 > I_k$) специфичные ограничения:

$$R_1 < \beta R_k; \quad R_2 < \beta R_k. \quad (18-1)$$

Если желательно иметь степени насыщения обоих транзисторов примерно одинаковыми, следует принять:

$$R_1 = R_2 = R.$$

Ниже будет показано, что неравенства (18-1) не должны быть сильными. Поэтому в большинстве случаев можно считать $R \gg R_k$.

Определим длительности полупериодов T_1 и T_2 (рис. 18-2). Для этого запишем выражение

$$u_{62}(t) = U_{62}(\infty) + [U_{62}(0) - U_{62}(\infty)]e^{-\frac{t}{\tau_1}},$$

где $\tau_1 = C_1 R_1$;

$$U_{62}(0) = U_{c1}(0) = E_k - I_{k0} R_k;$$

$$U_{62}(\infty) = U_{c1}(\infty) = -(E_k + I_{k0} R_1).$$

Полагая $U_{62}(T_1) = 0$, получаем длительность первого полупериода:

$$T_1 = \tau_1 \ln \frac{2E_k + I_{k0}(R_1 - R_k)}{E_k + I_{k0}R_1}.$$

Учитывая, что $R_1 \gg R_k$, приведем выражение для первого полупериода к следующему виду:

$$T_1 = C_1 R_1 \ln \frac{2 + \Phi}{1 + \Phi}. \quad (18-2a)$$

Аналогично для второго полупериода получим:

$$T_2 = C_2 R_2 \ln \frac{2 + \Phi}{1 + \Phi}, \quad (18-2b)$$

где

$$\Phi = \frac{I_{k0} R}{E_k} \quad (18-3)$$

является фактором теплового тока, характеризующим отношение последнего к насыщающему току базы.

Если тепловой ток $I_{к0}$ отсутствует (кремниевые транзисторы или низкая рабочая температура), то формулы (18-2) упрощаются:

$$T_1 = \tau_1 \ln 2 \approx 0,7C_1R_1; \quad (18-4a)$$

$$T_2 = \tau_2 \ln 2 \approx 0,7C_2R_2. \quad (18-4b)$$

В полностью симметричном мультивибраторе, в котором

$$R_1 = R_2 = R; \quad C_1 = C_2 = C,$$

длительности полупериодов будут одинаковыми и полный рабочий период

$$T = T_1 + T_2 = 2CR \ln \frac{2+\Phi}{1+\Phi}. \quad (18-5a)$$

В факторе Φ скрыта температурная зависимость полупериодов, а следовательно, и рабочей частоты. С увеличением температуры увеличивается ток $I_{к0}$ (т. е. фактор Φ), длительность полупериодов уменьшается, а рабочая частота растет. Поэтому величину Φ стараются делать малой. Тогда в формуле (18-5a) можно разложить логарифм в ряд с точностью до членов первого порядка и получить приближенное выражение

$$T \approx 2CR \left(\ln 2 - \frac{\Phi}{2} \right). \quad (18-5b)$$

Задавая приемлемым значением Φ , находят необходимую величину R из (18-3) [конечно, с учетом (18-1)], а затем величину C из формул (18-5).

Одним из способов повышения температурной стабильности является применение транзисторов с малым током $I_{к0}$, т. е. кремниевых. Однако, поскольку тепловые токи — не единственная причина неустойчивости, использование кремниевых транзисторов не всегда решает задачу стабилизации, особенно при пониженных температурах. Не менее важными причинами являются температурная зависимость напряжений на переходах насыщенного транзистора, а также зависимость $\beta(T)$. Как и в триггерах с эмиттерной связью, неустойчивость напряжений на переходах можно отразить с помощью э. д. с. $e = \epsilon \Delta T$ (см. § 17-3), добавляя их к начальному и конечному значениям напряжения на времязадающем конденсаторе. Тогда, как легко убедиться, к фактору Φ в формулах (18-5) добавятся еще и отношения вида $\epsilon \Delta T / E_k$, которые сохраняются при $I_{к0} = 0$. Зависимость $\beta(T)$ проявляется в изменении момента времени, соответствующего началу регенерации, поскольку этот момент определяется условием $\beta_1 \beta_2 \geq 1$. С ростом температуры значения β растут и регенерация наступает раньше, т. е. длительность полупериодов сокращается, а частота увеличивается.

Основными времязадающими элементами в транзисторном мультивибраторе являются емкости C_1 и C_2 , поскольку сопротивления R_1 и R_2 ограничены сверху условием насыщения (18-1),

а снизу либо временем рассасывания, либо (если одновременно уменьшать R_k) потребляемой мощностью.

Увеличение емкостей в принципе не ограничено, хотя сопровождается существенным удлинением отрицательного фронта, о чем подробнее говорится ниже. Уменьшение же емкостей (повышение рабочей частоты) ограничено условием равенства положительного и отрицательного фронтов, так же как в триггере [см. (16-33) и (16-46)]. Воспользуемся условием (16-46б) как более сильным и запишем его в следующем виде:

$$C_{\min} R_k \approx 3\tau_a. \quad (18-6)$$

Выражение (18-6) по аналогии с триггером позволяет разумно выбирать сопротивления R_k в начале расчета. А именно, поскольку стараются делать $C_{\min} \gg C_k$, сопротивления R_k должны быть достаточно малыми.

Кроме того, выражение (18-6) позволяет оценить максимальную частоту мультивибратора при заданном типе транзисторов или ориентировочно подобрать тип транзисторов по заданной частоте. Для этого положим $\phi = 0$ в формуле (18-5б) и подставим в нее величину $C_{\min}$ из (18-6). Тогда

$$T_{\min} \approx 4\tau_a \frac{R}{R_k}. \quad (18-7a)$$

Ниже будет показано, что сопротивление R должно быть по крайней мере в 3—4 раза больше R_k . С другой стороны, согласно (18-1) оно не может быть больше, чем βR_k . Разумным компромиссом, обеспечивающим степень насыщения $N \approx 2 \div 3$, является обычно соотношение

$$R = 10R_k.$$

При этом общее выражение (18-7a) переходит в более простое, удобное для ориентировочных оценок:

$$T_{\min} = \frac{1}{F_{\max}} \approx 40\tau_a. \quad (18-7б)$$

Например, при $\tau_a = 10$ нс получается $F_{\max} \approx 2,5$ МГц.

Спецификой мультивибратора, построенного на дрейфовых транзисторах, является то, что эмиттерный переход закрывшегося транзистора (T_2 на рис. 18-1) пробивается при небольшом напряжении $U_{\text{проб}} = 1 \div 2$ В (см. с. 241) и после этого быстро разряжает конденсатор C_1 . В результате начальное напряжение $U_{c1}(0)$ будет равно не $E_k - I_{k0}R_k \approx E_k$, как у бездрейфовых транзисторов, а напряжению $U_{\text{проб}} \ll E_k$. Соответствующее возрастание частоты можно, вообще говоря, скомпенсировать увеличением емкости, однако в процессе разряда ток через эмиттерный переход может быть слишком большим и транзистор может выйти из строя. Поэтому при использовании дрейфовых транзисторов обычно используют защитные диоды (см. рис. 15-22),

Регулирование частоты. Осу­ществлять плавную регулировку частоты с помощью переменных конденсаторов далеко не всегда возможно. Удобнее пользоваться переменными резисторами R , хотя при этом меняется степень насыщения транзисторов.

В схеме, показанной на рис. 18-3, для регулировки частоты используется переменное смещение. Такая схема позволяет менять частоту не только перемещением движка делителя напряжения, но и наложением сравнительно низкочастотного переменного напряжения на постоянное смещение. Очевидно, что таким путем может осуществляться частотно-импульсная модуляция. Покажем, к каким изменениям приводит питание резисторов $R_{1,2}$ от источника $E_{см} \neq E_k$. Для этого рассмотрим тот же полупериод T_1 , в течение которого транзистор T_1 насыщен, а транзистор T_2 заперт.

Прежде всего заметим, что коллекторный ток насыщения остается прежним и равным E_k/R_k , тогда как минимальный ток базы в насыщенном транзисторе уменьшается и равен теперь $E_{см}/R_1$. При этом условие насыщения приводит к соотношениям, отличным от (18-1):

$$R_1 < \beta \epsilon_{см} R_k; \quad R_2 < \beta \epsilon_{см} R_k, \quad (18-8)$$

где $\epsilon_{см} = E_{см}/E_k$ — относительное смещение. При $\epsilon_{см} < 1$ условия (18-8) требуют меньших значений времязадающих сопротивлений.

Перейдем к определению длительности периода. Конденсатор C_1 будет разряжаться от того же начального напряжения $E_k - I_{к0}R_k$, что и раньше, но конечным напряжением будет теперь не $-(E_k + I_{к0}R_1)$, а $-(E_{см} + I_{к0}R_1)$. Используя тот же метод, что и при выводе формул (18-2), легко получить величины T_1 и T_2 . Для полностью симметричного мультивибратора период выражается следующим образом:

$$T = 2CR \ln \frac{(1 + \epsilon_{см}) + \phi}{\epsilon_{см} + \phi}. \quad (18-9)$$

Сравнивая (18-9) и (18-5а), приходим к выводу, что уменьшение относительного смещения, т. е. напряжения $E_{см}$, увеличивает период и, значит, уменьшает рабочую частоту. При этом, конечно, подразумевается, что условия (18-8) выполняются для наименьшего значения $\epsilon_{см}$. Дифференцируя (18-9) по $\epsilon_{см}$, легко получить чувствительность регулировки (или модуляции) частоты и убедиться в том, что она растет с уменьшением смещения. Относительная чувствительность $\left(\frac{dF/F}{d\epsilon_{см}} \right)$ уменьшается с уменьшением $\epsilon_{см}$ и при $\epsilon_{см} = 0,5$ составляет около единицы.

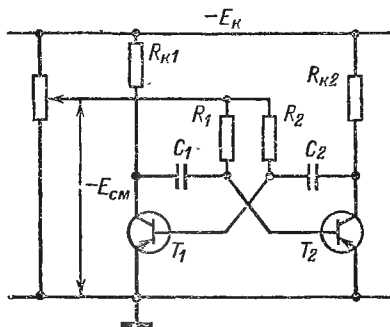


Рис. 18-3. Мультивибратор с регулируемым смещением.

С уменьшением напряжения смещения (рис. 18-8) все более затрудняется и в конце концов делается практически невозможным. В этом отношении наиболее показательным является случай $E_{см} = 0$, когда резисторы R просто заземлены. Ламповый аналог такой схемы, как известно, вполне работоспособен [162], хотя и обладает пониженной стабильностью частоты. Транзисторный мультивибратор с «заземленными» базами практически неработоспособен, так как он генерирует импульсы резко искаженной формы (рис. 18-4). Это объясняется тем, что открытый транзистор не может быть насыщен в течение всего полупериода, так как ток базы стремится к нулевому значению. На рис. 18-4 транзистор T_1 выходит из насыщения в некоторый момент $t_1 < T_1$ и далее работает в активном режиме. Поэтому последующее уменьшение базового тока вызывает пропорциональное уменьшение $I_{к1}$ и $U_{к1}$. Приращения $\Delta U_{к1}$ передаются через C_1 на базу второго транзистора и вызывают ускоренный спад $U_{б2}$. В результате описанных процессов не только изменяется длительность полупериода, но и резко искажаются отрицательные фронты выходных импульсов, особенно токов (рис. 18-4). Можно показать, что одновременно увеличивается длительность положительных фронтов, так как ток базы отпирающего транзистора оказывается значительно меньшим, чем в основной схеме (рис. 18-1).

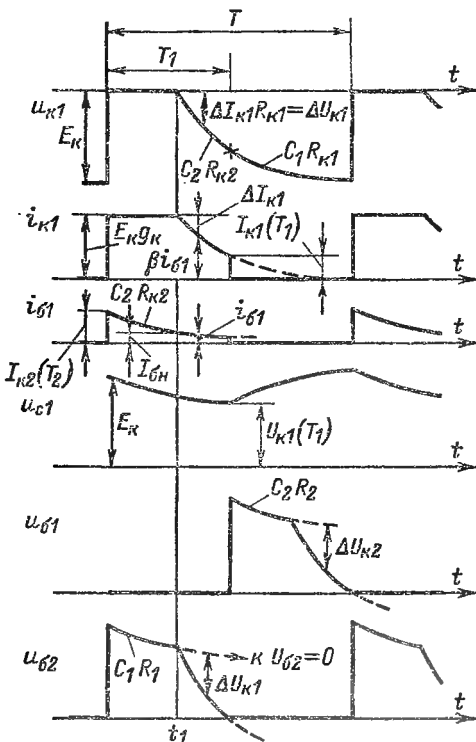


Рис. 18-4. Временные диаграммы работы мультивибратора с «заземленными» базами.

пульса и паузы. Для определенности будем считать T_1 длительностью импульса, а T_2 — длительностью паузы. Тогда *скважность* импульсов

$$Q = \frac{T_2}{T_1} + 1.$$

Будем рассматривать только случай $Q > 2$, так как случай $Q < 2$ означает просто изменение полярности импульса при сохранении всех закономерностей.

Главным препятствием на пути увеличения скважности является большая длительность отрицательного фронта (рис. 18-2). Последний имеет экспоненциальный характер и во время импульса T_1 формируется на коллекторе второго транзистора с постоянной

Скважность импульсов и отрицательный фронт.

Нередко требуется иметь разные полупериоды T_1 и T_2 или, что то же самое, разные длительности им-

времени $C_2 R_K$. На уровне $0,9 E_K$ длительность отрицательного фронта равна:

$$t_{\Phi_2} = 2,3 C_2 R_K.$$

Разделив эту величину на длительность импульса (18-4а), получим:

$$\frac{t_{\Phi_2}}{T_1} = 3,3 \frac{C_2 R_K}{C_1 R_1}.$$

Отсюда следует, что отрицательный фронт будет короче импульса при условии

$$\frac{R_1}{R_K} > 3,3 \frac{C_2}{C_1}. \quad (18-10)$$

Для того чтобы обеспечить одинаковую степень насыщения обоих транзисторов, сопротивления R_1 и R_2 выбирают одинаковыми. Тогда величины T_2 и T_1 будут пропорциональны соответствующим емкостям, а отношение C_2/C_1 будет равно $Q - 1$. Значит, как следует из (18-10), максимальная скважность зависит от величины R_1/R_K , а последняя согласно (18-1) всегда меньше β . Таким образом, максимальная скважность равна:

$$Q_{\text{макс}} \approx \frac{\beta}{3} + 1. \quad (18-11)$$

Например, при $\beta = 50$ получается $Q_{\text{макс}} \approx 18$. Практически с учетом конечной степени насыщения это значение будет по крайней мере в 2 раза меньше.

Из выражения (18-10) можно сделать еще один важный вывод. Учитывая, что минимальная скважность $Q_{\text{мин}} = 2$, а значит, минимальное отношение $C_2/C_1 = 1$, получаем условие

$$R_1 > 3,3 R_K. \quad (18-12)$$

Как видим, в нормально работающем мультивибраторе *время-задающие сопротивления не могут быть сколь угодно малыми* даже при соблюдении условий насыщения. Ограничение (18-12) было использовано выше при выводе соотношения (18-7б).

Чем сильнее выполняется неравенство (18-10), тем меньшую долю полупериода занимает время t_{Φ_2} , т. е. тем меньше *относительное* искажение отрицательного фронта. Однако абсолютное значение t_{Φ} при низкой рабочей частоте может быть весьма большим.

Одним из способов дополнительного укорочения отрицательного фронта является диодная фиксация коллекторных потенциалов на уровне E_{Φ} , меньшем напряжения E_K . Такая схема показана на рис. 18-5, а, а соответствующие временные диаграммы — на рис. 18-5, б. Из последних легко выразить время t_{Φ_2} на уровне $0,9 E_{\Phi}$:

$$t_{\Phi_2} = C_2 R_K \ln \frac{1}{1 - 0,9 \epsilon_{\Phi}}, \quad (18-13a)$$

где $\epsilon_{\Phi} = E_{\Phi}/E_K$ — относительный уровень фиксации.

При условии $\epsilon_\phi < 0,5$ можно разложить логарифм в ряд с точностью до первого члена и получить приближенное выражение:

$$t_{\phi 2} \approx 0,9 \epsilon_\phi C_2 R_K \quad (18-136)$$

Что касается длительности импульса T_1 , то она выражается прежней формулой (18-2а) с той лишь разницей, что в фактор теплового тока ϕ войдет напряжение E_ϕ вместо E_K .

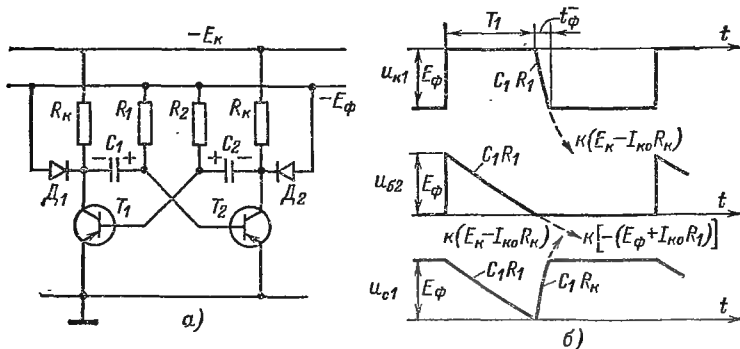


Рис. 18-5. Мультивибратор с фиксацией коллекторных потенциалов.
а — принципиальная схема; б — временные диаграммы.

Пренебрежем фактором ϕ в формуле (18-2а), разделим (18-136) на (18-2а) и положим это отношение меньшим единицы. Тогда получим неравенство, аналогичное неравенству (18-10):

$$\frac{R_1}{R_2} > \frac{0,9 \epsilon_\phi C_2}{\ln 2 C_1} \approx 1,3 \epsilon_\phi \frac{C_2}{C_1} \quad (18-14)$$

Отсюда легко найти ограничения на максимальную скважность и на отношение R_1/R_K , аналогичные ограничениям (18-11) и (18-12), но действительные только при $\epsilon_\phi < 0,5$:

$$Q_{\max} \approx 0,8 \frac{\beta}{\epsilon_\phi} + 1; \quad (18-15)$$

$$R_1 > 1,3 \epsilon_\phi R_K. \quad (18-16)$$

Как видим, схема с фиксацией обладает несомненными преимуществами в отношении длительности отрицательного фронта и максимальной скважности. Однако эти преимущества проявляются наиболее полно при условии $\epsilon_\phi \ll 1$, а это приводит к резкому снижению экономичности схемы. В самом деле, при расчете нужно выбирать э. д. с. E_ϕ равной заданному значению выходного напряжения. Тогда э. д. с. $E_K = E_\phi / \epsilon_\phi$ будет значительно больше E_ϕ и соответственно возрастут потери мощности.

Схема, показанная на рис. 18-6, а, позволяет получить длительность отрицательного фронта, практически равную длительности положительного фронта. Однако максимальная скважность импульсов в таком мультивибраторе оказывается меньше, чем в простой схеме. Соответствующие временные диаграммы показаны на рис. 18-6, б.

Во время стадии регенерации ток через диод, присоединенный к коллектору запирающегося транзистора, уменьшается и делается

равным нулю. Поэтому в конце стадии регенерации соответствующий конденсатор оказывается отключенным от запертого транзистора и последующий заряд его не влияет на величину коллекторного потенциала. Следовательно, длительность отрицательного фронта будет определяться только процессом запираания транзистора и, как в обычном ключе, может составлять всего несколько τ_a . Однако восстановление напряжения на конденсаторе происходит с постоянной времени CR_3 , которую нельзя сделать сколь угодно малой. В самом деле, при открытом состоянии транзистора резисторы R_3 и R_k соединены параллельно через открытый диод. Следовательно, если $R_3 \leq R_k$, то мощность, потребляемая схемой, существенно возрастает.

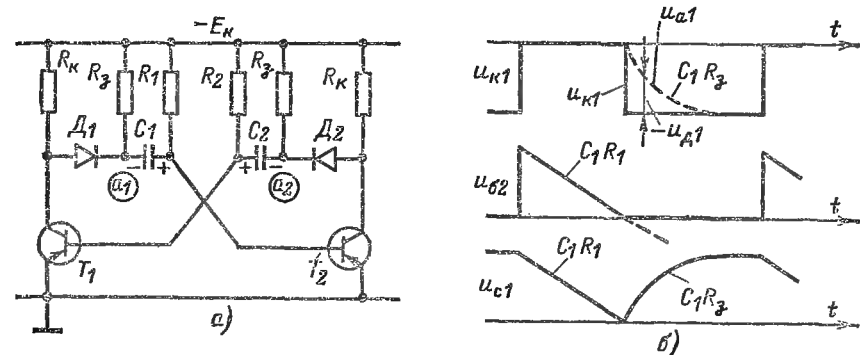


Рис. 18-6. Мультивибратор с блокировкой коллекторных потенциалов.

а — принципиальная схема; б — временные диаграммы.

Что касается скажности, то ее максимальное значение можно получить из условия (18-10), подставив R_3 вместо R_k (так как заряд конденсатора происходит через резистор R_3) и $R_1 = \beta (R_k \parallel R_3)$. Принятое значение R_1 соответствует границе насыщения открытого транзистора. Тогда максимальная скажность выразится следующим образом:

$$Q_{\text{макс}} \approx \frac{\beta}{3} \frac{R_k}{R_k + R_3} + 1.$$

В работе [169] для рассматриваемой схемы рекомендуется соотношение $R_3 \approx R_k$. При этом потребляемая мощность будет примерно вдвое больше, а максимальная скажность вдвое меньше, чем в простом мультивибраторе.

18-2. МУЛЬТИВИБРАТОР С РАЗРЯДНЫМ ТРИГГЕРОМ

Введение. Рассмотренные выше симметричные мультивибраторы имеют много общего с симметричными триггерами. Естественно, что на основе триггера с эмиттерной связью (см. гл. 17) можно построить другое семейство мультивибраторов — несимметричных. Для этого, например, в схеме на рис. 17-1 нужно исключить резистор R_4 , присоединить резистор R_2 к шинке E_K , а источник сигнала E_T заме-

снизить соответствующим постоянным смещением. Такой мультивибратор имеет много общего с одновибратором с эмиттерной связью, рассматриваемым в следующей главе. Возможны и другие варианты несимметричных мультивибраторов, которые связаны с исходным триггером не столь непосредственно [168, 169].

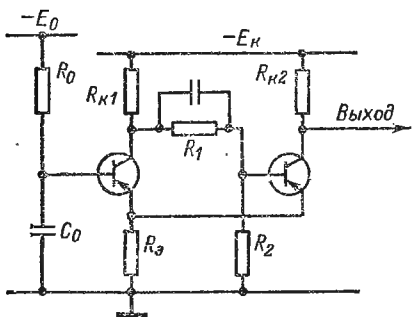


Рис. 18-7. Принципиальная схема мультивибратора с разрядным триггером.

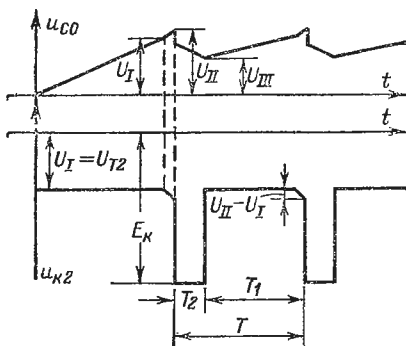


Рис. 18-8. Временные диаграммы напряжений на конденсаторе и на выходе мультивибратора (см. рис. 18-7).

Ниже исследуется схема, построенная по принципу так называемого «неоновых релаксатора» ([164], гл. 20), в которой триггер с эмиттерной связью играет роль разрядника для конденсатора, входящего в простейшую интегрирующую RC-цепочку (рис. 18-7). Эта схема может использоваться и в качестве генератора пилообразного напряжения, если выходом являются зажимы конденсатора C (см. § 21-3).

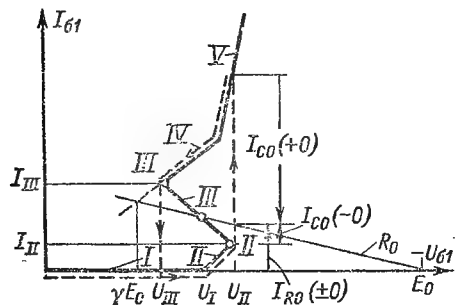


Рис. 18-9. Рабочий цикл мультивибратора (см. рис. 18-7), построенный по входной характеристике триггера.

Рабочий цикл. В первый момент после подключения источника питания напряжение на конденсаторе C_0 равно нулю и триггер находится в исходном состоянии (транзистор T_1 заперт, T_2 насыщен). Конденсатор заряжается через резистор R_0 (а также остаточным током базы запертого транзистора T_1) по экспоненциальному закону с постоянной времени $C_0 R_0$ (рис. 18-8). Такой процесс заряда нарушается при $U_{C_0} = U_I$, когда отпирается транзистор T_1 . Для анализа последующих процессов удобно использовать входную характеристику

триггера, изученную в § 17-2. Эта характеристика вместе с линией нагрузки R_0 воспроизведена на рис. 18-9. На участке II входное сопротивление триггера шунтирует конденсатор C_0 и ускоряет его заряд. Когда напряжение U_{C_0} достигнет значения U_{II} , триггер опрокинется. При этом рабочая точка окажется на участке V входной характеристики и входной ток резко возрастет. Как видно из рис. 18-9, ток через конденсатор изменяет знак и он начинает разряжаться. Разряд будет происходить по участкам V и IV, из которых первый имеет очень малое, а второй несколько большее сопротивление. Во всяком случае разряд будет происходить достаточно быстро до напряжения U_{III} , при котором триггер вер-

нется в исходное состояние, а конденсатор C_0 снова начнет медленно заряжаться.

Описанный рабочий цикл возможен в том случае, когда линия R_0 пересекает характеристику только в одной точке на «отрицательном» участке *III*. В противном случае либо процесс заряда закончится на участке *II* без срабатывания и я триггера (рис. 18-10, *а*), либо после срабатывания процесс разряда закончится на участке *IV* без отпущения (рис. 18-10, *б*). Таким образом, для режима автоколебаний необходимо соблюдение условия

$$\frac{E_0 - U_{III}}{I_{III}} < R_0 < \frac{E_0 - U_{II}}{I_{II}}, \tag{18-17}$$

которое легко получить из рис. 18-9. Сделанные замечания не специфичны для данной схемы, а являются общими при анализе устройств с неоднозначными (*S*-образными) вольт-амперными характеристиками.

Рабочая частота и скважность. Исходя из приведенного описания и рисунков, можно рассчитать длительность каждого этапа процесса в самом общем случае. Однако мы проиллюстрируем возможности схемы в упрощенном варианте, приняв $U_I \approx U_{II} \approx U_{IV}$. В этом случае (выкладки см. в предыдущих изданиях книги)

$$T_1 = C_0 R_0 \ln \frac{E_0 - U_{III}}{E_0 - U_{II}}; \tag{18-18}$$

$$T_2 = C_0 (R_0 \parallel R_{IV}) \ln \frac{U_{II} - \gamma E_0}{U_{III} - \gamma E_0}. \tag{18-19}$$

В выражении (18-19) всегда соблюдается неравенство $\gamma E_0 < U_{III}$; это легко показать на основании левой части неравенств (18-17).

Одним из преимуществ данного мультивибратора является возможность обеспечения большой скважности импульсов (т. е. большого отношения T_1/T_2); для этого необходимо сделать $R_0 \gg R_{IV}$. Чтобы не нарушать при этом правую часть неравенства (18-17), иногда приходится одновременно увеличивать напряжение E_0 .

Вторым преимуществом данного мультивибратора являются хорошая форма выходных импульсов и слабое влияние нагрузки на работу схемы — свойства, обусловленные применением триггера с эмиттерной связью.

К числу недостатков можно отнести необходимость дополнительного источника питания E_0 .

Сокращение длительности импульса (интервала T_2) требует уменьшения сопротивления R_{IV} , т. е. (поскольку $R_{IV} \sim R_3$) сопротивления R_3 . Сопротивление R_3 согласно (17-186) пропорционально сопротивлению $R_{к2}$, определяющему все параметры схемы. Значит, уменьшение интервала T_2 по существу означает применение низких сопротивлений с соответствующим понижением экономичности схемы.

Температурная стабильность интервалов T_1 и T_2 определяется стабильностью сопротивления R_{IV} [т. е. согласно (17-12) — стабильностью коэффициента β], а также стабильностью порогов срабатывания и отпущения U_{II} и U_{III} (см. § 17-3). Для германиевых транзисторов на длительность паузы T_1 оказывает влияние еще и ток $I_{к0}$ в базе запертого транзистора. Этот ток, которым мы пренебрегли при выводе (18-18), должен быть учтен добавлением слагаемого $I_{к0} R_0$ к ве-

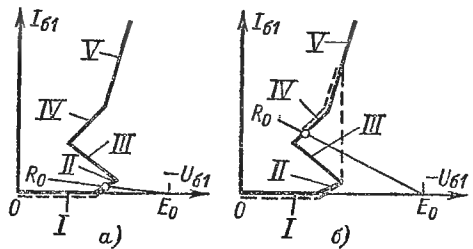


Рис. 18-10. Случаи неправильного выбора зарядного сопротивления R_0 и зарядной э. д. с. E_0 .

а — заряд конденсатора заканчивается на участке *II*; *б* — заряд конденсатора заканчивается на участке *IV*.

личине E_0 , стоящей под логарифмом. Для того чтобы влияние тока $I_{к0}$ не было большим, нужно соблюдать условие $E_0/R_0 \gg I_{к0}$.

Дополнительные сведения о влиянии $I_{к0}$ и других факторов на длительность паузы можно почерпнуть из § 20-2, посвященного аналогичным проблемам в блокинг-генераторах.

Глава девятнадцатая

ОДНОВИБРАТОРЫ

19-1. ОДНОВИБРАТОР С ЭМИТТЕРНОЙ СВЯЗЬЮ

Одновибратор — спусковая схема с самостоятельным возвращением в исходное состояние — сочетает в себе многие свойства триггеров и мультивибраторов¹. Наиболее распространен на практике одновибратор с эмиттерной связью, который мы рассмотрим более подробно.

Схема на рис. 19-1 является «производной» от триггера с эмиттерной связью. При изучении ее мы, как и в случае мультивибратора, будем считать изменения токов при переходе транзисторов из одного состояния в другое мгновенными. Иначе говоря, пренебрежем этапами рассасывания и регенерации в переходном процессе.

Рабочий цикл. В исходном состоянии транзистор T_1 заперт, а транзистор T_2 насыщен. Запирание транзистора T_1 обеспечивается делителем R_1, R_2 , который задает на базу первого транзистора потенциал U_0 меньший (по модулю), чем потенциал эмиттеров; последний равен потенциалу U_{T2} второго, насыщенного транзистора.

Подадим на вход отрицательный спусковой сигнал с амплитудой, превышающей запирающее напряжение $|U_{T2} - U_0|$. Под действием этого сигнала одновибратор опрокидывается во временно устойчивое состояние, когда транзистор T_2 заперт, а T_1 открыт и насыщен. Для простоты положим, что спусковой импульс очень короткий, и не будем учитывать влияния его на дальнейшее поведение схемы. Тогда переходные процессы будут такими, как показано на рис. 19-2.

Основным фактором, определяющим время выдержки, является перезаряд конденсатора C . В исходном состоянии конденсатор заряжен до напряжения $E_k = -I_{к0}R_{к1} - U_{T2}$. После опрокидывания он разряжается и напряжение на нем стремится к величине $-(E_k + I_{к0}R - U_{T1})$. Эту величину легко получить, мысленно

¹ Помимо термина «одновибратор», для схем такого типа применяются термины «запертый», «ждущий», «заторможенный» мультивибратор, «однотактный релаксатор» и некоторые другие. Термин «одновибратор» представляется наиболее правильным, так как он устанавливает определенную связь с термином «мультивибратор», исключает подчиненную роль по отношению к мультивибратору (это необходимо потому, что обе схемы равноправны), отражает однотактный режим работ и является одним из самых кратких.

убрав из схемы запертый транзистор T_2 и заменив его со стороны базы генератором тока $I_{к0}$. Когда напряжение U_c проходит через нуль, открывается второй транзистор, поскольку насыщенный транзистор T_1 считается точкой и можно принять $U_{бэ2} \approx U_c$. После отпираания транзистора T_2 транзистор T_1 быстро выходит из насыщения и начинается регенерация, в результате которой получается исходное состояние транзисторов. Однако состояние схемы в целом при этом отличается от исходного, так как напряжение на конденсаторе C после второго опрокидывания близко к нулю и лишь через некоторое время станет равным исходному значению. Соответственно потенциалы U_{T2} , $U_{к1}$ лишь постепенно приходят к начальному значению. Этот этап переходного про-

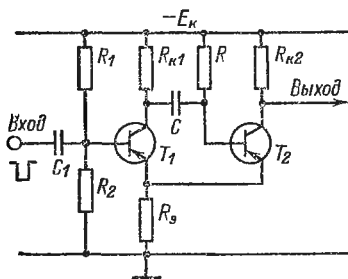


Рис. 19-1. Принципиальная схема одновибратора с эмиттерной связью.

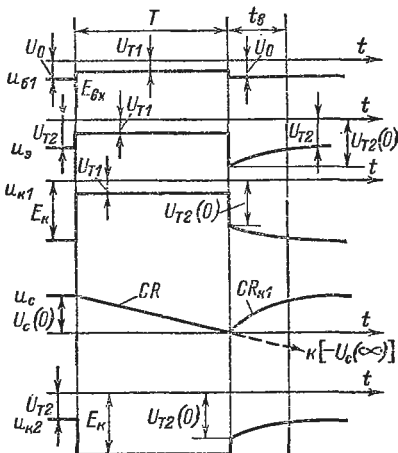


Рис. 19-2. Временные диаграммы напряжений в одновибраторе при коротком спусковом импульсе.

цесса, называемый *этапом восстановления*, играет значительную роль в работе одновибраторов. До полного восстановления схема при поступлении очередного импульса либо выдает уменьшенный и укороченный импульс, либо вообще не срабатывает.

Потенциал U_{T1} во время выдержки T меняется пренебрежимо мало, так как он определяется сравнительно низкоомными сопротивлениями (R_9 , R_2 и др.), на которых небольшой ток перезаряда конденсатора не дает заметного падения напряжения.

Статический расчет. Исходный потенциал второго транзистора определяется формулой узлового напряжения (15-11а):

$$U_{T2} = \frac{E_k (g_{к2} + g)}{g_{к2} + g + g_9}. \quad (19-1)$$

Значениями U_{T2} и $R_{к2}$ обычно задаются в начале расчета, как и в случае триггера с эмиттерной связью. При выборе U_{T2} руководствуются очевидным условием

$$U_{вык} = E_k - U_{T2}, \quad (19-2)$$

а при выборе $R_{к2}$ — соотношением (18-6).

Токи $I_{\beta 2}$ и $I_{\kappa 2}$ выражаются соотношениями

$$I_{\beta 2} = (E_{\kappa} - U_{T2}) g;$$

$$I_{\kappa 2} = (E_{\kappa} - U_{T2}) g_{\kappa 2};$$

подставляя их в условие насыщения $\beta_2 I_{\beta 2} > I_{\kappa 2}$, легко получить следующее ограничение на величину R :

$$R < \beta_2 R_{\kappa 2}. \quad (19-3)$$

Неравенство (19-3) не должно быть сильным во избежание глубокого насыщения. Поэтому, как правило, $R \gg R_{\kappa 2}$. Тогда, пренебрегая в формуле (19-1) проводимостью g , можно получить для сопротивления R_2 :

$$R_2 \approx \frac{U_{T2}}{E_{\kappa} - U_{T2}} R_{\kappa 2}. \quad (19-4)$$

Делитель напряжения R_1, R_2 рассчитывается из условия $|U_0| < |U_2| = |U_{T2}|$, где U_0 — потенциал $U_{\beta 1}$ в исходном состоянии.

Выразив потенциал U_0 с учетом тока $I_{\kappa 0}$ в базе закрытого транзистора T_1 , получим:

$$\frac{R_2}{R_1 + R_2} (E_{\kappa} + I_{\kappa 0} R_1) < U_{T2}. \quad (19-5)$$

Отсюда легко определить необходимое отношение R_2/R_1 .

Выражение (19-5) показывает, что с точки зрения температурной стабильности желательно иметь R_1 , а значит, и делитель в целом достаточно низкоомными. Для кремниевых транзисторов такое ограничение несущественно.

Нетрудно убедиться, что потенциал U_{T1} после срабатывания одновибратора несколько меньше U_0 , поскольку через делитель протекает ток базы. Полагая для простоты

$$U_{T1} = U_0 \approx E_{\kappa} \frac{R_2}{R_1 + R_2}$$

и используя условие насыщения $\alpha_1 I_{\beta 1} > I_{\kappa 1}$, можно получить приближенное выражение для выбора сопротивления $R_{\kappa 1}$:

$$R_{\kappa 1} > \frac{E_{\kappa} - U_0}{U_0} R_2 \quad (19-6)$$

(коэффициент α_1 положен равным единице). Неравенство (19-6) следует выполнять с достаточным запасом.

Время выдержки и его стабильность. Процессом, определяющим время выдержки, является перезаряд конденсатора C . Подставим в выражение для заряда емкости

$$u_c(t) = U_c(\infty) + [U_c(0) - U_c(\infty)] e^{-\frac{t}{\tau_1}}$$

начальное и конечное значения напряжения U_c :

$$U_c(0) = (E_{\kappa} - I_{\kappa 0} R_{\kappa 1}) - U_{T2};$$

$$U_c(\infty) = -[(E_{\kappa} + I_{\kappa 0} R) - U_{T1}].$$

Полагая $U_c(T) = 0$, легко найти время выдержки T в следующем общем виде:

$$T = \tau_1 \ln \frac{2E_{\kappa} - (U_{T1} + U_{T2}) + I_{\kappa 0} (R - R_{\kappa 1})}{E_{\kappa} - U_{T1} + I_{\kappa 0} R}. \quad (19-7a)$$

Здесь постоянная времени приблизительно равна:

$$\tau_1 \approx CR.$$

Более строго к величине R следовало бы добавить эквивалентное сопротивление $R_3 \parallel R_{к1} \parallel R_1 \parallel R_2$, подключенное к узловой точке T_1 .

Формулу (19-7а) целесообразно упростить, учитывая обычно выполняющиеся соотношения $R \gg R_k$ и $U_{T1} < U_{T2} \ll E_k$. Такое упрощение позволяет получить более наглядное выражение, по форме совпадающее с выражением для полупериода в мультивибраторе:

$$T \approx CR \ln \frac{2 + \phi}{1 + \phi}. \quad (19-7б)$$

Здесь, как и в мультивибраторе, ϕ — фактор теплового тока, определяемый формулой (18-3). Считая фактор ϕ малым, можно еще больше упростить выражение для T , разложив логарифм в ряд с точностью до членов первого порядка относительно ϕ . Тогда получим выражение, аналогичное (18-5б):

$$T \approx CR \left(\ln 2 - \frac{\phi}{2} \right). \quad (19-7в)$$

Задаввшись фактором ϕ , легко найти сопротивление R из определения (18-3), соблюдая, конечно, условие (19-3). После этого из выражений (19-7) находят необходимую емкость C .

Уменьшение времени выдержки достигается уменьшением емкости. Однако минимальное значение последней должно удовлетворять условию (18-6). Подставим $C_{\min}$ из (18-6) в (19-7) и положим $\phi = 0$. Тогда минимальное время выдержки будет определяться отношением R/R_k . Это отношение, как увидим ниже, невыгодно делать малым из-за роста времени восстановления. С другой стороны, согласно (19-3) оно не может превышать β_2 . Поэтому типичным значением можно считать $R/R_k = 10 \div 20$, как и в случае мультивибратора. Тогда связь между минимальным временем выдержки и постоянной времени τ_α принимает следующий вид:

$$T_{\min} = (20 \div 40) \tau_\alpha. \quad (19-8)$$

Это соотношение полезно для оценки возможностей данного типа транзистора или подбора необходимого типа. Для дрейфовых транзисторов такое соотношение действительно только тогда, когда есть возможность соответственно уменьшить R_k (см. § 16-5). В противном случае минимальное время выдержки будет в несколько раз, а нередко и на порядок больше, чем следует из (19-8).

Относительно температурной стабильности времени выдержки можно сделать те же замечания, что и в случае мультивибратора (см. § 18-1).

Время восстановления. Потенциал транзистора T_2 непосредственно после обратного опрокидывания легко найти, учитывая, что в этот момент $U_c = 0$ и, следовательно, $R_{к1}$ и R соединены

параллельно. Тогда из формулы (15-11а) получаем:

$$U_{T_2}(0) = \frac{E_k (g_{k2} + g + g_k)}{g_{k2} + g + g_3 + g_{k1}}. \quad (19-9)$$

Этот потенциал больше (по модулю), чем конечное значение, определяемое формулой (19-1). Одновибратор возвращается в исходное состояние по мере заряда конденсатора C . Заряд конденсатора C происходит по цепи, состоящей из сопротивления R_{k1} и параллельно соединенных сопротивлений, сходящихся в «узловую точку» T_2 . Обычно основное значение имеет сопротивление R_{k1} ; тогда постоянная времени восстановления выражается простым соотношением

$$\tau_2 = CR_{k1}.$$

С такой же постоянной времени потенциал $u_{T_2}(t)$ возвращается к исходному значению U_{T_2} :

$$u_{T_2}(t) = U_{T_2} + [U_{T_2}(0) - U_{T_2}]e^{-\frac{t}{\tau_2}}. \quad (19-10)$$

Уровень, на котором определяется время восстановления, как всегда, условен. Целесообразно исходить из того, чтобы в момент условного восстановления t_b потенциал $U_{T_2}(t_b)$ отличался от исходного значения на заданную величину ΔU_{T_2} . Тогда будут достаточно определенными возможные изменения чувствительности схемы и ее выходного напряжения. Полагая $U_{T_2}(t_b) = U_{T_2} + \Delta U_{T_2}$, легко найти:

$$t_b = \tau_2 \ln \frac{U_{T_2}(0) - U_{T_2}}{\Delta U_{T_2}}.$$

Если подставить сюда значения (19-1) и (19-9), получится довольно громоздкое выражение. Его можно упростить, полагая $R \gg R_3$, $R_{k1} = R_{k2}$ и используя выражение (19-4). Тогда после преобразований приведем время восстановления к следующему приближенному виду:

$$t_b \approx CR_{k1} \ln \left(\frac{U_{T_2}}{\Delta U_{T_2}} \frac{E_k - U_{T_2}}{E_k + U_{T_2}} \right). \quad (19-11)$$

Как видим, излишне большая величина сопротивления R_{k1} в данном случае вредна.

Малое время восстановления всегда желательно при разработке одновибраторов, так как при этом уменьшается допустимый интервал между входными импульсами. Полагая $\vartheta = 0$ и $U_{T_2} \ll E_k$, легко получить из (19-7) и (19-11) следующее соотношение между временем восстановления и временем выдержки:

$$\frac{t_b}{T} \approx \frac{R_k}{R} \frac{\ln(U_{T_2}/\Delta U_{T_2})}{\ln 2}.$$

При типичных значениях $\Delta U_{T_2} = (0,1 \div 0,01) U_{T_2}$ и $R = (10 \div 20) R_K$ получаем:

$$\frac{t_B}{T} = 0,15 \div 0,6.$$

Если необходимо сократить время восстановления, можно использовать схему с фиксацией коллекторного потенциала U_{K1} на уровне $E_\phi < E_K$. Этот способ известен из предыдущей главы и, так же как в мультивибраторах, приводит к увеличению потребляемой мощности. Если потенциал E_ϕ получается с помощью делителя, а не отдельного низкоомного источника, то фиксирующий диод можно использовать для подачи спусковых импульсов, как показано на рис. 19-3. Здесь фиксирующим является диод D_1 , а диод D_2 играет вспомогательную роль, разряжая конденсатор C_1 , если он зарядился за время входного импульса. Специфика такого «коллекторного» запуска рассматривается в следующем разделе.

Роль спускового импульса. Выше мы считали входной импульс коротким и не учитывали его влияния. Между тем такое влияние имеет место. Пусть, например, в схеме на рис. 19-1 амплитуда импульса заметно больше пороговой величины $|U_{T_2} - U_0|$, а его длительность сравнима со временем выдержки. Тогда работа одновибратора будет характеризоваться кривыми на рис. 19-4. Не вдаваясь в подробности анализа, можно отметить, что во время действия импульса предельное напряжение $U_c(\infty)$ уменьшается на величину $(E_{вх} + U_0) - U_{T_1}$, а начальное напряжение не меняется.

Следовательно, если бы импульс продолжался в течение всего времени выдержки, то перезаряд конденсатора происходил бы по кривой, которая на рис. 19-4 показана пунктиром. На самом деле в момент окончания импульса предельное значение $U_c(\infty)$ восстанавливается и кривая $u_c(t)$ претерпевает излом. В результате вторая часть перезаряда протекает быстрее, чем первая, но в целом время выдержки получается больше, чем при коротком импульсе. Таким образом, величину $t_{вх}$ желательно иметь малой по сравнению с T . В случае необходимости следует укорачивать входной сигнал. Для этого емкость конденсатора C_1 выбирают из условия

$$C_1 R_{вх} < t_{вх}. \quad (19-12)$$

Здесь $R_{вх} = R_1 \parallel R_2 \parallel R_{K1} \parallel R_3$ — результирующее сопротивление

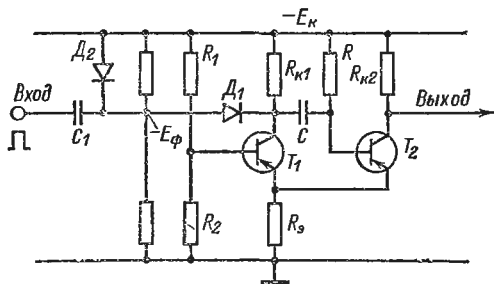


Рис. 19-3. Одновибратор с фиксацией коллекторного потенциала первого транзистора.

ние в «узловой точке» T_1 ; вместе с емкостью C_1 оно образует укорачивающую (дифференцирующую) RC -цепочку.

Очередной импульс, поступающий на вход одновибратора в интервале T , также замедляет перезаряд конденсатора C , если длительность $t_{вх}$ сравнима со временем выдержки. Кроме того, такой импульс частично просачивается на выход через емкости C и $C_{к2}$ и дает короткие отрицательные «пики» на коллекторе второго транзистора.

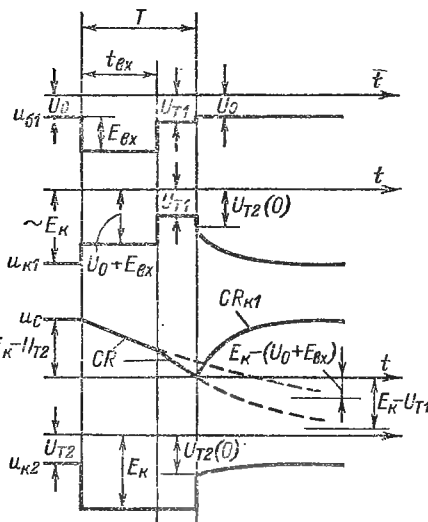
Схема с коллекторным запуском (см. рис. 19-3) в принципе мало отличается от схемы с базовым запуском, но все же имеет некоторые особенности. Прежде всего сигнал имеет в этом случае положительную полярность, так как он, поступая через

конденсатор C в «узловую точку» T_2 , должен уменьшить (по модулю) потенциал U_{T2} по крайней мере до значения U_0 , с тем чтобы открылся первый транзистор. Соответственно на выходе одновибратора в момент поступления сигнала получается положительный выброс, а не отрицательный, как при базовом запуске. Однако это обстоятельство существенно лишь при анализе фронтов. Значение $E_{вх}$ должно по-прежнему превышать $|U_{T2} - U_0|$, но запирающий импульс $\Delta I_{б2}$ обычно имеет большее значение и рассасывание происходит быстрее.

Рис. 19-4. Временные диаграммы напряжений в одновибраторе при большом и длинном спусковом импульсе.

Главная особенность коллекторного запуска состоит в том, что фиксирующий диод D_1 запирается в процессе переброса схемы. Это ясно из того, что после переброса потенциал его катода $U_{к1}$ равен U_{T1} , а это значение значительно меньше (по модулю), чем $E_{ф}$. Следовательно, в интервале выдержки источник сигналов отключен от одновибратора. Поэтому ни длительность импульса $E_{вх}$, ни поступление второго импульса в интервале T не оказывают влияния на время выдержки. Это, конечно, является достоинством схемы с коллекторным запуском, так же как и в случае симметричного триггера (см. § 16-5).

Анализируя одновибратор, мы подразумевали, что источник сигналов является генератором э. д. с., т. е. $R_r = 0$ или, точнее, $R_r \ll R_{вх}$. Под входным сопротивлением понимается та же величина, что и в условии (19-12). Если R_r имеет значительную вели-



чину, то замедляются процессы формирования фронтов и ухудшается чувствительность схемы, поскольку на вход поступает не все напряжение $E_{вх}$, а только часть его, определяемая соотношением сопротивлений $R_Г$ и $R_{вх}$.

19-2. ОДНОВИБРАТОР С РАЗРЯДНЫМ ТРИГГЕРОМ

Данный вариант одновибратора (рис. 19-5) по структуре идентичен соответствующему мультивибратору (рис. 18-7) и отличается от него лишь значениями E_0 и R_0 , а также наличием цепи запуска. Анализ этой схемы проведем с помощью входной характеристики триггера.

В § 18-2 было показано, что автоколебания возможны только тогда, когда пересечение входной характеристики и линии нагрузки R_0 имеет место в одной точке, на участке III (см. рис. 18-9). На рис. 18-10 были проиллюстрированы режимы, в которых автоколебания невозможны. В случае одновибратора именно эти режимы являются рабочими. Возьмем за основу тот из них, в котором ток через R_0 меньше (рис. 18-10, а).

На рис. 19-6 показан цикл работы одновибратора после подачи спускового сигнала в виде импульса от генератора э. д. с. При этом емкость C_1 будем считать достаточно большой, чтобы пренебречь изменениями ее заряда за время входного импульса и далее, вплоть до восстановления схемы. Такую большую емкость можно принять за источник э. д. с. с напряжением U_{A0} . Тогда рабочий цикл можно описать следующим образом.

До поступления сигнала рабочей точкой одновибратора является точка 1 на участке II. Потенциал точки А (см. рис. 19-5) имеет значение U_{A0} , показанное на рис. 19-6; линия нагрузки R_2 , проведенная из абсциссы U_{A0} , пересекает характеристику в той же точке 1.

При подаче отрицательного входного импульса $E_{вх}$ линия нагрузки R_2 смещается вправо по оси абсцисс на величину $E_{вх}$. Если э. д. с. $E_{вх}$ достаточно велика, точка пересечения попадает на участок V (рис. 19-6, а) или IV (рис. 19-6, б). Напряжение на конденсаторе C в первый момент остается неизменным, но под действием появившегося зарядного тока оно начинает увеличиваться. В точке II происходит скачкообразное изменение тока через конденсатор, причем в зависимости от сигнала ток I_C может либо сохранять свое направление (рис. 19-6, а), либо изменить знак (рис. 19-6, б). В первом случае (сигнал $E_{вх1}$) конденсатор стремится разрядиться до напряжения в устойчивой точке 2. Во втором случае (сигнал $E_{вх2}$) конденсатор стремится разрядиться до напряжения в устойчивой точке 3. Для достижения устойчивых точек нужно некоторое время. Поэтому если длительность входного импульса недостаточно велика, то напряжение U_C не достигнет значения U_2 или U_3 и к концу импульса останется близким к значению $U_4 \approx U_{II}$ (рис. 19-6, в). При этом, конечно, подразумевается, что длительность $t_{вх}$ больше, чем время, необходимое для перехода рабочей точки из исходного положения 1 в критическую точку II и скачка на участок V вольт-амперной характеристики.

В момент окончания входного сигнала потенциал точки А (см. рис. 19-5) снова делается равным U_{A0} . При этом, линия R_2 возвращается в исходное положение.

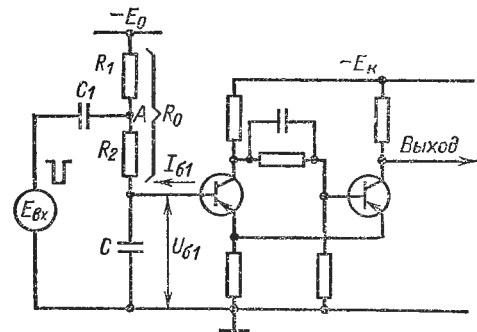


Рис. 19-5. Одновибратор с разрядным триггером.

ние и ток I_c протекает в разрядном направлении. Напряжения U_c уменьшается и рабочая точка движется влево по участку IV (рис. 19-6, б) или по участку V, а потом по участку IV (рис. 19-6, а). В точке III происходит скачок тока в конденсаторе и ток приобретает з а р я д н о е направление. В результате конденсатор заряжается до исходного напряжения в точке I. На этом рабочий цикл заканчивается. Из проведенного рассмотрения можно сделать ряд полезных выводов.

Очевидно, что выходной импульс (запертое состояние транзистора T_2) соответствует пребыванию и движению рабочей точки на участках V и IV. По-

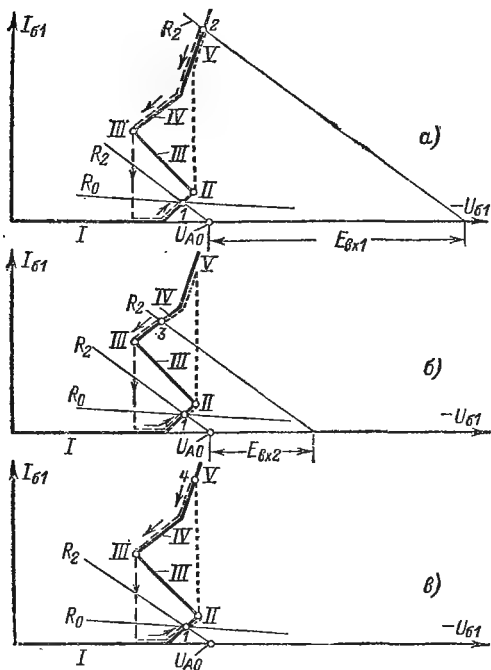


Рис. 19-6. Цикл работы одновибратора с разрядным триггером при двух значениях длинного сигнала (а, б) и при тех же значениях короткого сигнала (в).

этому время выдержки должно зависеть от положения рабочей точки в момент окончания входного сигнала. Для того чтобы время выдержки не зависело от амплитуды и длительности спускового сигнала, время $t_{вх}$ должно быть настолько малым, чтобы конденсатор C после скачка не успел существенно разрядиться или зарядиться (рис. 19-6, в). Тогда движение рабочей точки влево будет начинаться всегда от одного и того же места (на рис. 19-6, в — от точки 4, соответствующей напряжению U_{II}).

При длинном спусковом сигнале рабочая точка, вообще говоря, успевает переместиться в сторону устойчивой точки 2 или 3, а иногда и достигнуть этих точек (рис. 19-6, а и б). Тогда время выдержки будет зависеть и от амплитуды, и от длительности сигнала — факт, отмеченный на с. 567, где анализ проводился другим методом.

Время восстановления соответствует движению рабочей точки по участкам I и II после обратного скачка. На участке I первый транзистор заперт и, значит, заряд конденсатора происходит через R_2 с посто-

янной времени CR_2 . Уменьшение времени восстановления путем уменьшения R_2 имеет предел, поскольку линия R_2 не должна пересекать участок III характеристики. Кроме того, при малом значении R_2 сильно нагружается источник сигнала.

Спусковой сигнал такой величины, при которой пересечение линии R_2 с входной характеристикой происходит не на участке V или IV, а на участке III, тоже вызывает срабатывание одновибратора. Однако, если такой сигнал будет длинным (в пределе — бесконечно длинным), схема перейдет в режим мультивибратора. Минимальный сигнал, необходимый для запуска, должен обеспечить пересечение линии R_2 с участком III немного левее точки II.

В заключение отметим, что скачки токов в точках II и III на самом деле не мгновенны, а определяются переходными процессами в триггере.

БЛОКИНГ-ГЕНЕРАТОР

Блокинг-генератор представляет собой релаксационную схему с трансформаторной обратной связью, выполненную на одном активном элементе. По сравнению с мультивибраторами и одно-вибраторами блокинг-генераторы позволяют получить большую скважность и меньшую длительность импульсов, а также осуществить трансформаторную связь с нагрузкой.

В транзисторных блокинг-генераторах транзистор может быть включен как по схеме ОБ, так и по схеме ОЭ, кроме того, используется несколько вариантов включения времязадающего конденсатора, нагрузки и т. п. [168, 169, 162]. Мы подробно проанализируем только основную схему (рис. 20-1), поскольку полученные выводы действительно в общем и для других вариантов.

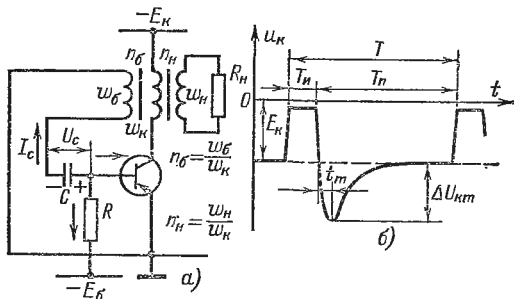


Рис. 20-1. Простейшая схема блокинг-генератора (а) и временная диаграмма коллекторного напряжения (б).

20-1. ОБЩЕЕ ОПИСАНИЕ

Схема блокинг-генератора, работающего в режиме автоколебаний, показана на рис. 20-1. Пусть сначала на конденсаторе C напряжение имеет положительную полярность (показанную на рис. 20-1, а) и достаточно большое значение. Тогда транзистор находится в запертом состоянии и конденсатор разряжается через R и вторичную обмотку трансформатора. Индуктивностью последней можно пренебречь, так как скорость изменения тока на этом этапе невелика. Строго говоря, разряд происходит не только через R , но и через цепь базы; это будет учтено при последующем анализе. Когда напряжение на конденсаторе, стремящееся к отрицательному уровню $-E_б$, падает до нуля, транзистор отпирается и начинается второй этап переходного процесса. На этом этапе возрастающий коллекторный ток наводит на вторичной обмотке трансформатора э. д. с. отрицательной полярности, которая способствует форсированному отпираанию транзистора. Процесс развивается лавинообразно вплоть до насыщения транзистора, когда последний утрачивает усилительные свойства и положительная обратная связь нарушается. Коллекторное напряжение в течение второго этапа падает от величины $-E_к$ практически до нуля. На третьем этапе происходит рассасывание накопленных в базе носителей. При этом

напряжение $U_{кз}$ остается близким к нулю, т. е. формируется вершина импульса. Через некоторое время заряд неосновных носителей в базе уменьшается до такой величины, при которой транзистор выходит из насыщения. После этого наступает этап запира-ния, в течение которого коллекторный ток лавинообразно падает до нуля, а коллекторное напряжение снова достигает значения $-E_k$. Затем получается выброс, обусловленный рассеянием магнитной энергии, накопленной в сердечнике трансформатора. По оконча-нии выброса схема возвращается в ис-ходное состояние и рассмотренный цикл повторяется.

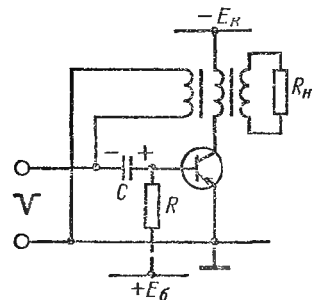


Рис. 20-2. Схема ждущего блокинг-генератора.

То положительное напряжение на конденсаторе, которое мы предположи-ли в начале цикла, получается в про-цессе формирования вершины, когда че-рез конденсатор протекает большой ба-зовый ток. Фронты импульса обычно бывают короткими, и напряжение U_c за время фронта практически не меняется.

Ж д у щ и й режим блокинг-генератора обеспечивается включением запирающего напряжения в цепь базы последовательно с резистором R (рис. 20-2). В этом случае до подачи отрицательного спускового импульса транзистор будет на-ходиться в запертом состоянии, а напряжение на конденсаторе будет равно напряжению смещения E_k . После подачи спускового импульса начнется описанный выше цикл, в конце которого вос-становится исходное состояние.

Поскольку в ждущем режиме длительность паузы задается периодом спусковых импульсов, сопротивление R выбирают доста-точно малым, с тем чтобы разряд конденсатора заканчивался до при-хода очередного импульса.

Ниже анализируются отдельные этапы переходного процесса в режиме автоколебаний.

20-2. ИНТЕРВАЛ МЕЖДУ ИМПУЛЬСАМИ

Интервал между двумя импульсами (время паузы) определя-ется временем разряда конденсатора от некоторого максимального напряжения $U_{см}$ до напряжения отпираия $U_{с0}$, близкого к нулю¹.

¹ Напряжение отпираия — величина, разумеется, условная (см. рис. 15-3). В данном случае под $U_{с0}$ можно понимать то напряжение $U_{эб}$, при котором коэф-фициент β , зависящий от тока (см. § 4-6), становится достаточным для лавинообраз-ного процесса, характеризующего фронт импульса. Обычно напряжение $U_{с0}$ соответствует коллекторным токам порядка 10—100 мкА и составляет 50—100 мВ для германиевых и 200—300 мВ для кремниевых транзисторов. Учет напряжения $U_{с0}$ не носит принципиального характера, поэтому, как и в случае мультивибра-торов, мы им в дальнейшем пренебрежем, считая, что транзистор отпирается при $U_c = 0$.

Транзистор на этом этапе заперт, но в цепи базы протекает некоторый остаточный ток $I_0 \sim [см., например, (15-1в)]$. Кроме того, имеет место ток через сопротивления утечки $R_{y.к}$ и $R_{y.э}$, шунтирующие запертые переходы транзистора. На рис. 20-3 показана эквивалентная схема разряда конденсатора, исходя из которой можно записать уравнение разряда в следующем виде:

$$C \frac{dU_c}{dt} = I_c = -\frac{U_c + E_6}{R} - \frac{U_c}{R_{y.э}} - \frac{U_c + E_k}{R_{y.к}} - I_0. \quad (20-1)$$

Решая уравнение при начальном условии $U_c(0) = U_{cm}$ и полагая затем $U_c = 0$, находим время паузы:

$$T_n = \tau_n \ln \frac{U_{cm} + E_6 + \left(I_0 + \frac{E_k}{R_{y.к}} - \frac{E_6}{R_{y.к} \parallel R_{y.э}} \right) R^*}{E_6 + \left(I_0 + \frac{E_k}{R_{y.к}} - \frac{E_6}{R_{y.к} \parallel R_{y.э}} \right) R^*}, \quad (20-2)$$

где $\tau_n = CR^*$ — эквивалентная постоянная времени разряда; $R^* = R \parallel R_{y.к} \parallel R_{y.э}$ — эквивалентное разрядное сопротивление.

Формула (20-2) имеет ту же структуру, что и формула (18-9) для мультивибратора, и является обобщением последней. На практике резистор R присоединяют к шине источника питания; тогда $E_6 = E_k$ и сходство между обеими формулами делается еще более очевидным.

Поскольку время паузы T_n пропорционально постоянной времени τ_n , желательно, чтобы последняя по возможности не зависела от нестабильных сопротивлений утечек. Для этого нужно стремиться к выполнению условия

$$R \ll R_{y.к} \parallel R_{y.э}, \quad (20-3)$$

которое обычно соответствует значениям R , не превышающим 50 кОм.

Если положить $R_{y.к} = R_{y.э} = \infty$ и $I_0 = 0$, то выражение (20-2) упрощается и характеризует длительность паузы в идеализированной схеме, в которой отсутствуют дестабилизирующие факторы:

$$T_n = RC \ln \left(\frac{U_{cm}}{E_6} + 1 \right). \quad (20-4)$$

В отсутствие смещения ($E_6 = 0$) стабильность блокинг-генератора оказывается весьма низкой, так как при этом время паузы согласно (20-2) существенно зависит от обратного тока и сопротивлений утечек. Практически такая схема, как и схема мультивибратора без смещения (рис. 18-4), неработоспособна.

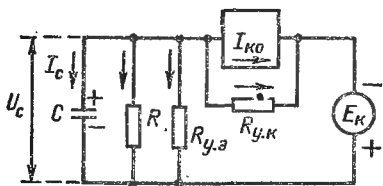


Рис. 20-3. Эквивалентная схема для интервала паузы.

Начальное напряжение U_{cm} определяется изменением базового потенциала на этапе формирования вершины импульса. Действительно, в интервале T_n (рис. 20-1, б) напряжение на коллекторной обмотке близко к E_k , так как транзистор насыщен и $U_{кэ} \approx 0$. Соответственно напряжение на базовой обмотке близко к $n_6 E_k$. В начале вершины все это напряжение падает на входе транзистора, поскольку за время фронта напряжение U_c почти не меняется и остается близким к нулю. В дальнейшем напряжение $n_6 E_k$ перераспределяется между конденсатором и входом транзистора: по мере заряда конденсатора напряжение U_c возрастает, стремясь к $n_6 E_k$, а напряжение U_b уменьшается, стремясь к нулю. Если этот процесс успевает закончиться, то

$$U_{cm} \approx n_6 E_k. \quad (20-5)$$

В противном случае напряжение U_{cm} будет значительно меньше. Малые значения U_{cm} затрудняют реализацию необходимого времени паузы: приходится увеличивать R в ущерб стабильности [см. (20-3)].

Ниже будет показано, что оптимальный коэффициент трансформации обычно составляет десятые доли единицы. Поэтому напряжение U_{cm} , как правило, не превышает нескольких вольт.

В случае дрейфовых транзисторов с их малым пробивным напряжением эмиттерного перехода интервал T_n имеет следующую специфику. По окончании импульса, когда транзистор начинает запирается, потенциал базы повышается только до значения $U_{проб}$, после чего конденсатор C быстро разряжается через пробившийся p - n переход и напряжение на нем падает от U_{cm} до $U_{проб}$. Импульс

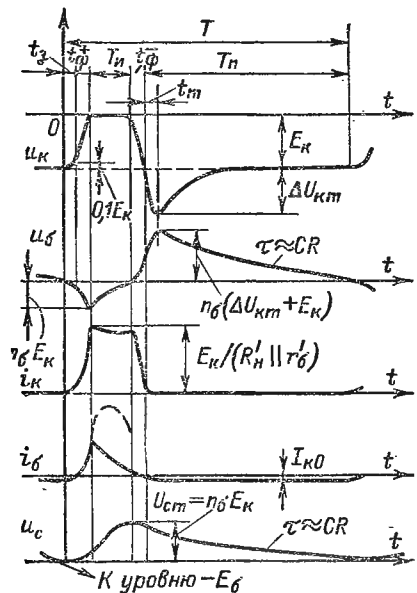


Рис. 20-4. Полные временные диаграммы блокинг-генератора со стабилизирующим смещением.

разрядного тока ограничивается сопротивлением r_6 . В остальном формирование паузы протекает так же, как описано выше, но в формулах (20-2) и (20-4) вместо напряжения U_{cm} должно стоять меньшее значение $U_{проб}$. Поскольку малые начальные напряжения нежелательны (см. выше), последовательно с базой (до резистора R) включают защитный диод (рис. 15-22), который ограничивает разрядный ток конденсатора даже в случае пробоя эмиттерного перехода. Обратный ток диода и его сопротивление утечки будут играть роль соответственно тока I_0 и сопротивлений $R_{у.к.}$, $R_{у.э}$ в формуле (20-2). Прямое падение напряжения на диоде во время формирования вершины приводит к некоторому уменьшению U_{cm} .

На рис. 20-4 показаны временные диаграммы, которые иллюстрируют работу блокинг-генератора как во время паузы, так и на последующих этапах.

Наличие трансформатора делает анализ фронтов в блокинг-генераторе еще более сложной задачей, чем в триггерах или мультивибраторах. Тем не менее такой анализ, хотя бы в упрощенном виде, необходим как для оценки длительности фронтов, так и для вывода соотношений, обеспечивающих оптимальную работу схемы.

На рис. 20-5 показана линейная эквивалентная схема коллекторной цепи на этапе формирования фронтов. Здесь трансформатор считается идеальным, т. е. не учитываются индуктивность обмотки L , индуктивность рассеяния L_s , межвитковая емкость C_L и сопротивление потерь в сердечнике R_L . Эти элементы включают в схему при более строгом анализе.

Сопротивление $R'_H = R_H/n_n^2$ есть сопротивление нагрузки, пересчитанное к коллекторной цепи. В частном, но распространенном случае, когда нагрузка R_H подключена к коллектору не через трансформатор, а через разделительный конденсатор, следует считать

$n_n = 1$. При трансформаторной связи с нагрузкой коэффициент трансформации выбирается из условия $n_n = U_H/E_K$, где U_H — желательное значение выходного напряжения. Таким образом, пересчитанное сопротивление R'_H в дальнейшем рассматривается как заданная величина.

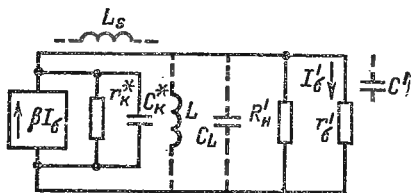


Рис. 20-5. Эквивалентная схема для интервала формирования фронтов.

На схеме показаны также пересчитанные к коллекторной обмотке параметры базовой цепи $r'_b = r_b/n_b^2$ и $C' = Cn_b^2$. Строго говоря, вместо r_b следовало бы использовать сопротивление R_{bx} . Однако, во-первых, сопротивление R_{bx} нелинейно (поскольку $r_b \sim 1/I_b$), а во-вторых, основная часть фронта формируется при достаточно больших токах I_b (десятки миллиампер), когда вполне приемлемо условие $r_b \approx 0$. Оба эти аргумента говорят в пользу приближения $R_{bx} \approx r_b$, которое использовано на эквивалентной схеме [см. также (15-15)]. Что касается емкости C , то она показана пунктиром, так как за время фронта приращение ΔU_c обычно несущественно, т. е. емкость можно считать бесконечно большой — «короткозамкнутой» по переменным составляющим.

Инерционность транзистора отражена, как обычно, операторным коэффициентом усиления $\beta(s)$ и операторным импедансом $Z_K(s)$ (параллельное соединение r_K^* и $1/sC_K^*$, см. § 4-7). С учетом обоих инерционных факторов приращения коллекторного тока будут определяться эквивалентным коэффициентом усиления (7-35):

$$\beta_{oe}(s) = \frac{\beta_{oe}}{1 + s\tau_{oe}},$$

где

$$\begin{aligned}\beta_{oe} &= \beta \gamma_k^*, \\ \tau_{oe} &= \tau \gamma_k^* + \tau_k (1 - \gamma_k^*), \\ \gamma_k^* &= \frac{r_k^*}{r_k^* + R_H' \| r_6'}.\end{aligned}$$

Согласно рис. 20-5 приращения $\Delta I_6'$ составляют некоторую долю от приращений ΔI_k , определяемую токораспределением между ветвями r_6' и R_H' :

$$\Delta I_6' (s) = \gamma_H [\beta_{oe} (s) \Delta I_6 (s)], \quad (20-6)$$

где

$$\gamma_H = \frac{R_H'}{R_H' + r_6'}.$$

Приращение истинного (непересчитанного) тока базы складывается из начального тока I_{60} , соответствующего напряжению U_{c0} , т. е. началу регенерации (см. сноску на с. 572), и приращения $\Delta I_6'$, пересчитанного в базовую цепь:

$$\Delta I_6 (s) = I_{60} + \frac{\Delta I_6' (s)}{n_6}.$$

Подставляя значение $\Delta I_6 (s)$ в (20-6) и решая получающееся уравнение относительно $\Delta I_6' (s)$, нетрудно привести изображение тока к следующему виду:

$$\Delta I_6' (s) = \frac{\beta_{oe} \gamma_H I_{60}}{(\beta_{oe} \gamma_H - 1) (s \tau^* - 1)}, \quad (20-7)$$

где $\beta_{oe}' = \beta_{oe}/n_6$ — приведенный эквивалентный коэффициент усиления (с учетом трансформации);

$$\tau^* = \frac{\tau_{oe}}{\beta_{oe} \gamma_H - 1} \quad (20-8)$$

— эквивалентная постоянная времени, характеризующая фронт импульса.

Оригиналом полученного изображения $\Delta I_6' (s)$ является нарастающая экспоненциальная функция. Если положить $\beta_{oe}' \gamma_H \gg 1$, что обычно выполняется, получим:

$$\Delta I_6' (t) = n_6 I_{60} \left(e^{\frac{t}{\tau^*}} - 1 \right). \quad (20-9)$$

Показатель экспоненты будет положительным и, следовательно, процесс будет лавинообразным только при условии $\tau^* > 0$. Учитывая выражения для τ^* и β_{oe}' , запишем это условие — *условие регенерации* — в виде

$$\beta' \gamma_k^* \gamma_H > 1 \quad (20-10a)$$

или, подставляя значения γ_k^* , γ_n , β' ,

$$\beta > n_6 \left(1 + \frac{r_6'}{R_n} + \frac{r_6'}{r_k^*} \right). \quad (20-10)$$

Правая часть неравенства (20-10) представляет собой тот минимальный коэффициент усиления, который соответствует значениям I_{60} и U_{c0} (см. сноску на с. 572).

Если обе части (20-9) умножить на $n_6 r_6'$, получим переходную характеристику приращений коллекторного напряжения:

$$\Delta u_k(t) = I_{60} r_6' \left(e^{\frac{t}{\tau^*}} - 1 \right). \quad (20-11)$$

Из (20-11), полагая $\Delta U_k(t_1) = 0,1 E_k$ и $\Delta U_k(t_2) = E_k$, находим время положительного фронта¹:

$$t_{\Phi}^* = 2,3 \tau^*.$$

Раскрывая значение постоянной времени τ^* и входящих в нее коэффициентов и полагая $\beta'_{oe} \gamma_n \gg 1$, $\tau/\beta_0 = \tau_a$, можно представить время фронта в следующем виде:

$$t_{\Phi}^* = 2,3 n_6 [\tau_a (1 + a_n) + C_k r_6'], \quad (20-12)$$

где $a_n = r_6'/R_n'$ — коэффициент нагрузки. Например, если $n_6 = 0,33$; $\tau_a = 10$ нс; $C_k = 10$ пФ; $r_6 = 100$ Ом и $R_n' = 200$ Ом, то $t_{\Phi}^* \approx 50$ нс.

Поскольку $r_6' = r_6/n_6^2$, легко заметить, что время нарастания имеет минимальное значение при некотором оптимальном коэффициенте трансформации. Продифференцировав по n_6 правую часть (20-12) и приравняв производную нулю, найдем:

$$n_{6, \text{опт}} = \sqrt{\frac{r_6}{R_n'} + \frac{r_6 C_k}{\tau_a}}. \quad (20-13)$$

Как видим, оптимальное значение n_6 помимо параметров самого транзистора зависит от приведенной нагрузки.

В случае холостого хода ($R_n' = \infty$)

$$n_{6, \text{опт}} = \sqrt{\frac{r_6 C_k}{\tau_a}}. \quad (20-14)$$

Например, если $r_6 = 100$ Ом; $C_k = 10$ пФ и $\tau_a = 10$ нс, то $n_{6, \text{опт}} \approx 0,33$. При значении R_n' , сравнимом с τ_a/C_k , значение $n_{6, \text{опт}}$ увеличивается. Минимум функции $t_{\Phi}^*(n_6)$ выражен довольно слабо, поэтому необязательно точно выдерживать расчетное значение $n_{6, \text{опт}}$. Во всех практических случаях n_6 получается меньше единицы.

Анализ фронта был проведен без учета индуктивности коллекторной обмотки L и времязадающей емкости C , т. е. считалось $L = \infty$ и $C = \infty$. Оценим

¹ Процесс, предшествующий моменту $\Delta U_k = 0,1 E_k$, будем считать задержкой (интервал t_z на рис. 20-4). Длительность задержки нередко превышает длительность фронта. Однако нелинейность параметров в этом интервале делает его анализ крайне трудным.

условия, при которых эти предположения практически оправданы, т. е. параметры L и C мало влияют на длительность фронта. Для этого положим постоянные времени $L/(R'_H \parallel r'_\delta)$ и $C'r'_\delta$ достаточно большими по сравнению со временем нарастания (20-12). Тогда получаются следующие неравенства, которые следует выполнять при проектировании блокинг-генератора:

$$L \gg 2,3n_6 r'_\delta [\tau_a + C_K (R'_H \parallel r'_\delta)]; \quad (20-15)$$

$$C \gg \frac{3,3}{n_6} \left(\frac{\tau_a}{R'_H \parallel r'_\delta} + C_K \right). \quad (20-16)$$

Если эти неравенства не соблюдаются, то время фронта будет больше, чем следует из (20-12).

Что касается индуктивности рассеяния трансформаторных обмоток L_s , то она должна быть достаточно малой, чтобы не вносить заметной инерционности в процесс нарастания базового тока. Действительно, из рис. 20-5 видно, что наличие индуктивности L_s препятствует изменениям тока в цепи $R'_H \parallel r'_\delta$. Если оценить эту инерционность с помощью постоянной времени $L_s/(R'_H \parallel r'_\delta)$ и положить последнюю много меньшей t_Φ^+ , то получится условие, обратное условию (20-15):

$$L_s \ll 2,3n_6 r'_\delta [\tau_a + C_K (R'_H \parallel r'_\delta)]. \quad (20-17)$$

Считая, что в каждом из неравенств (20-15) и (20-17) левая часть отличается от правой в 3—5 раз, приходим к выводу, что коэффициент рассеяния $\sigma_s = L_s/L$ должен составлять 5—10%.

Задний, отрицательный фронт импульса на участке $0 \div (-E_K)$ формируется примерно в тех же условиях, что и передний. Поэтому будем считать его длительность близкой к t_Φ^+ . Некоторое различие обоих фронтов обусловлено тем, что во время отрицательного фронта сопротивление базы модулировано, т. е. имеет меньшее значение.

20-4. ВЕРШИНА ИМПУЛЬСА

Будем считать длительностью импульса T_n время, в течение которого транзистор насыщен и коллекторное напряжение близко к нулю (вершина импульса на рис. 20-4). Временем переднего (положительного) фронта пренебрежем, т. е. положим, что переход транзистора из закрытого состояния в насыщенное осуществляется мгновенно. Кроме того, пренебрежем сопротивлением участка коллектор — эмиттер, которое при достаточном насыщении обычно не превышает 1—10 Ом. Сопротивление участка база — эмиттер будем считать равным r_δ .

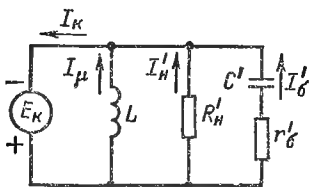


Рис. 20-6. Эквивалентная схема для интервала формирования вершины.

Пересчитывая к коллекторной обмотке величины R'_H , r_δ и C , получаем эквивалентную схему, показанную на рис. 20-6, где L — индуктивность коллекторной обмотки.

Поскольку передний фронт импульса принят бесконечно коротким, напряжение на конденсаторе не меняется за время фронта и остается равным напряжению отпирания транзистора в конце

паузы. Таким образом, в начале вершины имеем $U_c(0) \approx U_{c0}$; положим $U_c(0) = 0$. Тогда начальный ток насыщения $I_K(0) = E_K / (R'_H \parallel r'_6)$.

Из рис. 20-6 видно, что базовый ток меняется независимо от токов в других ветвях схемы, а коллекторный ток является суммой токов во всех трех ветвях. Поэтому изображения токов I'_6 и I_K запишутся следующим образом:

$$I'_6(s) = \frac{E_K}{r'_6 + 1/sC'}; \quad (20-18a)$$

$$I_K(s) = E_K \left(\frac{1}{R'_H} + \frac{1}{sL} \right) + \frac{E_K}{r'_6 + 1/sC'}. \quad (20-18б)$$

Оригиналы этих изображений после некоторых преобразований приводятся к виду

$$i'_6(t) = \frac{E_K}{r'_6} e^{-\frac{t}{\tau_c}}; \quad (20-19)$$

$$i_K(t) = E_K \left[\frac{1}{R'_H} + \frac{1}{r'_6} \left(\frac{t}{\tau_L} + e^{-\frac{t}{\tau_c}} \right) \right], \quad (20-20)$$

где

$$\tau_c = r'_6 C' = r_6 C;$$

$$\tau_L = L / r'_6.$$

Выражения (20-19) и (20-20) показывают, что ток базы монотонно уменьшается, а ток коллектора в зависимости от соотношения параметров может меняться по-разному. Если $\tau_L < \tau_c$, то ток I_K монотонно возрастает, так как доминирует линейный член t/τ_L . Если $\tau_L > \tau_c$, то ток I_K сначала уменьшается (доминирует экспоненциальный член), достигает минимума и затем непрерывно увеличивается. Именно этот типичный случай проиллюстрирован на рис. 20-4.

Коллекторный ток во время формирования вершины не должен превышать допустимого значения в импульсном режиме $I_{K, \text{доп}}$ (см. § 15-7). Поскольку согласно (20-20) увеличение тока во время импульса обусловлено линейным членом, положим для простоты $e^{-t/\tau_c} = 1$. Для этого наиболее неблагоприятного случая из (20-20) легко найти величину $I_K(T_H)$. Полагая $I_K(T_H) \leq I_{K, \text{доп}}$, получаем неравенство

$$L \geq \frac{T_H}{\frac{I_{K, \text{доп}}}{E_K} - \frac{1}{R'_H \parallel r'_6}}, \quad (20-21)$$

выполнение которого обеспечивает нормальный режим транзистора ($I_K \leq I_{K, \text{доп}}$) во всем интервале T_H . Если необходимая индуктивность L оказывается слишком большой, можно выбрать транзистор с большим значением $I_{K, \text{доп}}$. В неравенство (20-21) входит коэффициент трансформации n_6 , который оценивается по формуле (20-13).

Переходя к расчету длительности импульса, заметим, что конец интервала T_H соответствует выходу транзистора из насыще-

ния, поэтому $Q(T_n) = Q_{гр}$, где $Q_{гр}$ — граничный заряд, определяемый формулой (15-27). Функцию $Q(t)$ можно найти, решив уравнение заряда при известном токе базы (20-19). Что касается граничного значения $Q_{гр}$, то оно в нашем случае меняется со временем, так как ток насыщения описывается выражением (20-20)

Подставив (20-20) в (15-27) в качестве тока $I_{к.н}$, получим:

$$Q_{гр} = \frac{E_k \tau}{\beta} \left[\frac{1}{R'_n} + \frac{1}{r'_6} \left(\frac{1}{\tau_L} + e^{-\frac{t}{\tau_c}} \right) \right]. \quad (20-22a)$$

Для определения функции $Q(t)$ подставим в операторное уравнение (15-28a) изображение (20-18a) (разделив последнее на n_6 , чтобы оно было изображением реального, непересчитанного тока базы) и начальный заряд в базе, равный:

$$Q(0) = Q_{гр}(0) = \frac{\tau}{\beta} E_k / (R'_n \parallel r'_6).$$

Тогда изображение заряда примет следующий вид¹:

$$Q(s) = \frac{E_k (s\tau\tau_c)}{n_6 r'_6 (1+s\tau)(1+s\tau_c)} + \frac{E_k s\tau^2}{\beta (R'_n \parallel r'_6) (1+s\tau)}.$$

Этому изображению соответствует оригинал

$$Q(t) = \frac{E_k}{n_6 r'_6} \frac{e^{-\frac{t}{\tau}} - e^{-\frac{t}{\tau_c}}}{\frac{1}{\tau_c} - \frac{1}{\tau}} + \frac{E_k \tau}{\beta (R'_n \parallel r'_6)} e^{-\frac{t}{\tau}}. \quad (20-22б)$$

На рис. 20-7 показаны возможные варианты кривых $Q(t)$ и $Q_{гр}(t)$, пересечение которых определяет время T_n . Как видно, пересечение имеет место только при условии $(dQ/dt)_0 > (dQ_{гр}/dt)_0$. Это вполне естественно, так как в начале вершины транзистор находится на границе насыщения и в дальнейшем конечная степень насыщения получится только в том случае, если заряд в базе нарастает быстрее (или спадает медленнее), чем граничное значение $Q_{гр}$. Дифференцируя (20-22a) и (20-22б) и полагая $t = 0$, запишем *необходимое условие формирования вершины* в следующем виде:

$$\beta' > (1 + a_n) + \tau \left(\frac{1}{\tau_L} - \frac{1}{\tau_c} \right). \quad (20-23)$$

Здесь $\beta' = \beta/n_6$ — приведенный коэффициент усиления базового тока (с учетом трансформации); $a_n = r'_6/R'_n$ — коэффициент нагрузки ($a_n = 0$ при холостом ходе).

Из (20-23) видно, что индуктивность коллекторной обмотки должна быть достаточно большой, а емкость в базовой цепи —

¹ Здесь и далее постоянная времени τ соответствует режиму насыщения (см. §15-4); индекс «н» для простоты опущен. Некоторые сведения об этой постоянной времени применительно к блокинг-генератору можно найти в [170].

достаточно малой. В целом условия (20-23) являются более слабым, чем (20-106), так как последний член (20-23) можно всегда сделать меньше, чем r'_6/r_K^* .

Приравнявая (20-22a) и (20-22б), получаем для определения длительности импульса трансцендентное уравнение, которое после преобразований можно записать в форме

$$\frac{1}{\beta'} \left[a_n (e^{\vartheta_n} - 1) + \frac{\vartheta_n}{\vartheta_L} e^{\vartheta_n} \right] = \left(\frac{\vartheta_c}{\vartheta_c - 1} - \frac{1}{\beta'} \right) \left(e^{\vartheta_n \frac{\vartheta_c - 1}{\vartheta_c}} - 1 \right), \quad (20-24)$$

где $\vartheta_c = \tau_c/\tau$; $\vartheta_L = \tau_L/\tau$; $\vartheta_n = T_n/\tau$.

В общем случае решение уравнения (20-24) кропотливо. Однако это решение можно облегчить, если учесть, что на практике задача ставится следующим образом: известны параметры транзистора (τ , β , r_6), коэффициент трансформации n_6 и нагрузка R'_n ; требуется

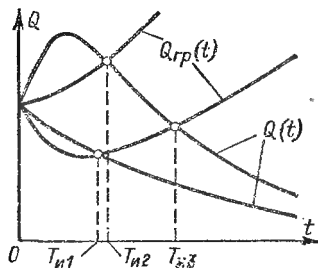


Рис. 20-7. Возможные виды функций заряда в базе Q и граничного заряда $Q_{гр}$.

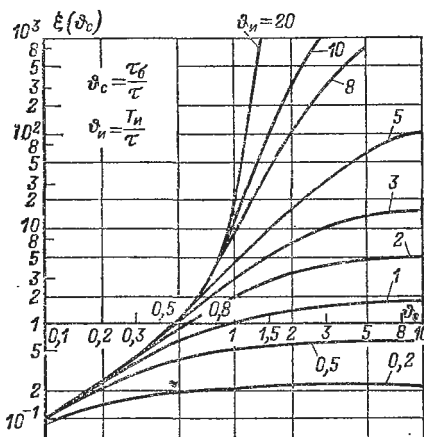


Рис. 20-8. Семейство характеристик функции $\xi(\vartheta_c; \vartheta_n)$, используемой при расчете длительности импульса.

найти значения величин ϑ_c и ϑ_L (т. е. по существу значения C и L), обеспечивающие заданное время T_n . Для решения этой задачи целесообразно преобразовать уравнение (20-24). Пренебрежем величиной $1/\beta'$ в правой части и введем обозначение

$$\xi(\vartheta_c; \vartheta_n) = \frac{\vartheta_c}{\vartheta_c - 1} \left(e^{\frac{\vartheta_n}{\vartheta_c}(\vartheta_c - 1)} - 1 \right). \quad (20-25)$$

Кроме того, введем обозначения для двух величин в левой части уравнения (20-24):

$$A_1 = \frac{a_n}{\beta'} (e^{\vartheta_n} - 1); \quad (20-26a)$$

$$A_2 = \frac{\vartheta_n}{\beta'} e^{\vartheta_n}. \quad (20-26б)$$

С учетом введенных обозначений уравнение (20-24) приводится к следующей форме, удобной для расчетов:

$$\xi(\vartheta_c; \vartheta_n) = A_1(\vartheta_n) + A_2(\vartheta_n)/\vartheta_L. \quad (20-27)$$

По исходным значениям ϑ_n , a_n , β' определяются коэффициенты A_1 и A_2 . После этого, задаваясь одним из параметров (ϑ_c или ϑ_L) и используя кривые на рис. 20-8, находим второй параметр. Например, если $a_n = 5$; $\beta' = 100$ и $\vartheta_n = 3$, то $A_1 \approx 1$ и $A_2 \approx 0,6$. Полагая $\vartheta_L = 0,5$, получаем $\xi \approx 2,2$ и $\vartheta_c \approx 0,8$, а полагая $\vartheta_c = 2$, получаем $\xi \approx 7$ и $\vartheta_L \approx 0,1$.

Заметим, что заданное значение ϑ_n можно реализовать только при достаточно больших значениях ϑ_c и ϑ_L . Так, если в предыдущем примере положить $\vartheta_n = 3$ и $\vartheta_c = 0,2$, то $\xi \approx 0,3$ и необходимое значение ϑ_L будет отрицательным, а если положить $\vartheta_L = 0,01$, то $\xi = 61$ и это значение при $\vartheta_n = 3$ нельзя получить ни при каких значениях ϑ_c . Ограничения на параметры ϑ_c и ϑ_L уже отмечались выше в связи с неравенством (20-23), которое соответствует реализации импульса с нулевой длительностью ($\vartheta_n = 0$).

В последующем анализе будут использоваться упрощенные выражения для функции ξ ; они получаются при достаточно больших или достаточно малых значениях ϑ_c и ϑ_n . А именно, если $\vartheta_c > 4$ и $\vartheta_n > 3$, то выражение (20-25) приводится к виду

$$\xi \approx e^{\vartheta_n} e^{-\vartheta_n/\vartheta_c}; \quad (20-28a)$$

а если $\vartheta_c < 0,2$ и $\vartheta_n > 2$, то¹

$$\xi \approx \vartheta_c. \quad (20-28b)$$

Условие (20-23) показывает, что в принципе блокинг-генератор сохраняет работоспособность при наличии только одной реактивности: C или L . Случай $C = \infty$, когда исчезает понятие не за-висимой паузы, используется в преобразователях постоянного напряжения (см. § 22-2, рис. 22-5).

Часто желательно, чтобы сопротивление нагрузки (если оно меняется) не оказывало влияния на длительность импульса. Для этого необходимо выполнить условие $A_1 \ll A_2/\vartheta_L$ [см. уравнение (20-27)], откуда легко получается неравенство

$$L \ll T_n R'_n \frac{e^{\vartheta_n}}{e^{\vartheta_n} - 1} \approx \tau R'_n (1 + \vartheta_n). \quad (20-29)$$

Здесь под R'_n понимается минимальное сопротивление нагрузки. Случай квазихолостого хода, характеризуемый неравенством (20-29), распространен на практике. Поэтому рассмотрим его подроб-

¹ Еще два варианта неравенств $\vartheta_c > 4$ и $\vartheta_n < 0,5$ или $\vartheta_c < 0,2$ и $\vartheta_n < 0,5$ приводят к одному и тому же приближению: $\xi \approx \vartheta_n$. Можно показать, что в этих случаях условие (20-23) выполняется очень слабо, т. е. указанные варианты практически неприменимы.

нее. Из условия $\xi = A_2/\phi_L$ получаем необходимое значение индуктивности:

$$L = \frac{r_6 T_n e^{\phi_n}}{\beta n_6 \xi}. \quad (20-30a)$$

В случае длинных импульсов и большой постоянной времени в цепи базы подставляем (20-28a); тогда после преобразований

$$L \approx \frac{r_6 T_n e^{T_n/\tau_c}}{\beta n_6}. \quad (20-30б)$$

Например, если $r_6 = 50$ Ом; $\tau = 5$ мкс; $T_n = 50$ мкс; $\beta = 50$; $C = 1,5$ мкФ; $n_6 = 0,2$, то $L \approx 50$ мкГ.

В случае малой постоянной времени в цепи базы и достаточно длинных импульсов в формулу (20-30a) подставляем (20-28в) и после преобразования получаем:

$$L \approx \frac{\tau_a T_n}{C n_6} \quad (20-30в)$$

(при выводе положено $\tau/\beta \approx \tau_a$). Например, если $\tau_a = 0,1$ мкс; $T_n = 2$ мкс; $n_6 = 0,2$; $C = 0,01$ мкФ, то $L \approx 100$ мкГ.

Формулы (20-30) приводят к важному выводу, который можно получить также из общего выражения (20-24): *время импульса находится в прямой зависимости от индуктивности L и емкости C* . Это понятно и из физических соображений, поскольку бесконечно большие значения L и C означают стационарное насыщение транзистора, т. е. бесконечно длинный импульс. В случае конечной (достаточно низкоомной) нагрузки, когда неравенство (20-29) не имеет места, длительность импульса уменьшается, т. е. зависимость T_n от R_n тоже прямая. Действительно, учитывая в уравнении (20-27) член A_1 , мы должны заменить значение ξ в формуле (20-30a) значением $\xi = A_1$. Следовательно, уменьшение R_n сопровождается как бы уменьшением ξ ; поэтому при неизменном значении L должно уменьшиться значение T_n .

До сих пор считалось, что ток базы начинает спадать в самом начале вершины [см. (20-19)]. Однако при достаточно больших времязадающих емкостях (обычно при $C > 0,1$ мкФ) наблюдается существенное возрастание базового тока в течение некоторого времени и только после этого ток начинает уменьшаться (показано пунктиром на рис. 20-4). Такое явление [170] объясняется модуляцией сопротивления r_6 накапливающимися в базе носителями (см. рис. 15-7). В самом деле, в момент $t = 0$ заряд сосредоточен в активной области базы, так как насыщение транзистора происходит в течение очень короткого переднего фронта. Затем дырки за время порядка τ распространяются в пассивную область базы и сопротивление r_6 уменьшается. Если напряжение $n_6 E_k = U_c$, действующее в цепи базы, уменьшается с постоянной времени τ_c большей, чем τ , то базовый ток будет возрастать до тех пор, пока r_6 достигнет установившегося значения. Получающийся выброс базового тока может в 2—3 раза превышать начальное значение $I_6(0)$.

Если рассчитывать емкость C , исходя из немодулированного начального значения r_{60} , то получающееся возрастание базового тока затянёт первый этап формирования вершины. Зато последующий этап, когда ток I_6 уменьшается, будет

протекать более быстро, так как постоянная времени $r_6 C$ будет меньше расчетной из-за модуляции сопротивления. Поэтому в целом время импульса определяется некоторой усредненной величиной $r_{6, \text{ср}}$. Аналитическое определение этой величины затруднено. Поэтому при формировании длинных импульсов ($T_n > \tau$) целесообразно включать последовательно с базой в цепи и постоянный резистор с сопротивлением R_6 , равным примерно r_{60} , с тем чтобы при расчете можно было не считаться с изменением суммарного сопротивления $R_6 + r_6$. При включении внешнего резистора R_6 получаются дополнительные преимущества: снижается мощность, расходуемая в цепи базы, и уменьшается влияние разброса величины r_{60} у разных транзисторов. Вводя резистор R_6 , нужно, разумеется, учитывать его сопротивление при расчете длительности импульса

20-5. ВЫБРОС НАПРЯЖЕНИЯ

Если бы индуктивность L была бесконечно большой и за время формирования вершины (а тем более за время фронтов) ток намагничивания не менялся, то в обмотке не запасалось бы никакой магнитной энергии и выброс не имел бы места. Чтобы оценить выброс, нужно определить ток намагничивания I_μ в конце интервала T_n , пользуясь эквивалентной схемой на рис. 20-6. Из схемы видно, что ток в ветви

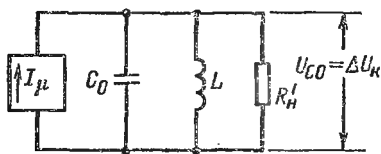


Рис. 20-9. Эквивалентная схема для интервала формирования выброса.

L нарастает независимо от остальных ветвей по закону $L \frac{dI_\mu}{dt} = E_k$, т. е. линейно. Поэтому в конце вершины ток будет равен:

$$I_\mu(T_n) = I_\mu = \frac{E_k T_n}{L}. \quad (20-31)$$

Считая задний фронт достаточно коротким, можно пренебречь изменением этого тока за время t_ϕ . Тогда в момент полного записывания транзистора получится эквивалентная схема, показанная на рис. 20-9. Здесь сопротивление R'_H , строго говоря, должно включать в себя также пересчитанное времязадающее сопротивление $R' = R/n_\beta^2$, подключенное параллельно R'_H ; однако эта поправка обычно незначительна. Суммарная емкость C_0 включает в себя пересчитанные емкости эмиттерного и коллекторного переходов¹:

$$C_0 = n_\beta^2 C_s + (1 + n_\beta) C_k. \quad (20-32)$$

В общем случае к этой емкости следует добавить пересчитанную емкость нагрузки.

Начальный ток I_μ , определяемый формулой (20-31), вызывает заряд емкости C_0 и соответственно повышение коллекторного напряжения. Величина $U_{C_0} = \Delta U_k$ найдется из операторного изображения

$$\Delta U_k(s) = I_c(s) \frac{1}{sC_0} = \left(I_\mu \frac{sL \parallel R'_H}{sL \parallel R'_H + \frac{1}{sC_0}} \right) \frac{1}{sC_0}.$$

¹ Емкость C_k , как видно из схемы на рис. 20-1, а, находится под напряжением $U_k - U_3 = U_k + n_\beta U_k$ (если считать $C = \infty$). Отсюда в формуле (20-32) появляется множитель $(1 + n_\beta)$.

Это изображение легко приводится к следующему простейшему виду:

$$\Delta U_{\kappa}(s) = I_{\mu} R'_{\text{н}} \frac{s\tau_{\text{в}}}{1 + s\tau_{\text{в}} + s^2 LC_0}, \quad (20-33)$$

где $\tau_{\text{в}} = L/R'_{\text{н}}$ — постоянная времени выброса.

Переходная функция в зависимости от соотношения параметров может иметь аperiодический или колебательный характер. Колебательный режим нежелателен, так как положительные выбросы, трансформируясь в базовую цепь в виде отрицательных импульсов, могут вызвать преждевременное отпирание транзистора и тем самым изменение рабочей частоты. В ждущем блокинг-генераторе положительные выбросы могут вызвать срабатывание схемы в отсутствие очередного сигнала. Учитывая сказанное, рассмотрим только аperiодический режим и прежде всего найдем критерий такого режима. Для этого приравняем нулю дискриминант знаменателя в формуле (20-33). Отсюда получим обычное условие критического режима в колебательном контуре:

$$\sqrt{\frac{L}{C_0}} = 2R_{\text{н}}. \quad (20-34)$$

Условием аperiодического режима будет неравенство

$$\sqrt{\frac{L}{C_0}} > 2R_{\text{н}}. \quad (20-35)$$

Это условие накладывает ограничение на значение приведенной нагрузки и, в частности, делает невозможным режим холостого хода

Оценим величину и длительность отрицательного выброса в критическом режиме. Заменив в изображении (20-33) коэффициент LC_0 на $\tau_{\text{в}}^2/4$ [согласно (20-34)], получим:

$$\Delta U_{\kappa}(s) = I_{\mu} R'_{\text{н}} \frac{s\tau_{\text{в}}}{\left(1 + s\frac{\tau_{\text{в}}}{2}\right)^2}.$$

Оригиналом такого изображения будет функция

$$\Delta U_{\kappa}(t) = 2I_{\mu} R_{\text{н}} \left(\frac{2t}{\tau_{\text{в}}} e^{-\frac{2t}{\tau_{\text{в}}}} \right).$$

Легко убедиться, что эта функция имеет максимальное значение, которое с учетом (20-31) равно:

$$\Delta U_{\kappa m} = \frac{2}{e} I_{\mu} R_{\text{н}} \approx 0,75 E_{\kappa} \frac{T_{\text{и}}}{\tau_{\text{в}}}. \quad (20-36)$$

Максимум имеет место в момент

$$t_m = \frac{\tau_B}{2} = \frac{L}{2R_B}. \quad (20-37)$$

Пусть, например, $L = 2 \text{ мГ}$; $R'_B = 1 \text{ кОм}$ и $I_\mu = 10 \text{ мА}$; тогда $t_m = 1 \text{ мкс}$ и $\Delta U_{км} \approx 7,5 \text{ В}$. Как видим, выброс имеет значительную величину; при других параметрах схемы он может быть еще больше.

Длительность выброса, учитывая его экспоненциальный спад, можно условно оценить временем

$$t_B = 2,5\tau_B, \quad (20-38)$$

которому соответствует спад выброса примерно до 3% максимального значения (20-36). При использованных выше параметрах получим $t_B \approx 5 \text{ мкс}$.

В случае аperiodического режима выражения для величин $\Delta U_{км}$ и t_m получаются громоздкими [170]. Анализ их приводит к выводу, что в «идеальном» аperiodическом режиме [когда неравенство (20-35) выполняется очень сильно] амплитуда выброса превышает величину (20-36) примерно на 30%, момент максимума $t_m \approx 0$, т. е. почти совпадает с моментом окончания заднего фронта, а длительность выброса составляет $(3 \div 4) \tau_B$.

Отрицательный выброс в коллекторной цепи трансформируется в базовую цепь в виде положительного импульса. При этом напряжение между коллектором и базой превышает значение $-E_K$:

$$U_{\text{кб. макс}} \approx -(E_K + \Delta U_{км})(1 + n_6). \quad (20-39)$$

Разумеется, значение $U_{\text{кб. макс}}$ должно быть меньше допустимого напряжения, определяемого пробоем коллекторного перехода.

Напряжение между эмиттером и базой в момент максимального выброса равно:

$$U_{\text{эб. макс}} \approx -n_6(E_K + \Delta U_{км}). \quad (20-40)$$

Это значение также не должно превышать допустимого напряжения, что практически нельзя обеспечить в случае дрейфовых транзисторов, у которых пробивное напряжение эмиттер — база обычно лежит в пределах 1—2 В. Например, в приведенном ранее расчете получилось $\Delta U_{км} = 7,5 \text{ В}$. Значит, при $E_K = 10 \text{ В}$ и $n_6 = 0,2$ имеем $U_{\text{эб. макс}} \approx -4 \text{ В}$, что превышает напряжение пробоя для многих транзисторов. Если в цепь базы включен защитный диод (см. рис. 15-22), то такое перенапряжение несущественно. Если же диода нет или если выброс вообще нежелателен, то коллекторную обмотку трансформатора шунтируют диодом или цепочкой диод — резистор. При обратной (по сравнению с этапом формирования вершины) полярности напряжения на коллекторной обмотке диод проводит и результирующим нагрузочным сопротивлением в формуле (20-36) будет $R'_B \parallel R_d \ll R'_B$.

ГЕНЕРАТОРЫ ПИЛООБРАЗНОГО НАПРЯЖЕНИЯ

21-1. ВВЕДЕНИЕ

Пилообразным напряжением называют одиночные или периодические импульсы такой формы, как показано на рис. 21-1. Поскольку пилообразное напряжение широко используется в электронных осциллоскопах для линейной развертки луча, генераторы пилообразного напряжения называют иногда *генераторами развертки*, хотя этот термин не отражает их гораздо более широкого применения. Из области разверток заимствованы названия двух основных частей пилообразного импульса: *прямой ход* (главный, почти линейный участок $T_{пр}$) и *сбратный ход* (сравнительно короткий участок $T_{обр}$, форма которого обычно несущественна). Еще один термин — *генераторы линейно изменяющегося напряжения* (ГЛИН) — правильно отражает основную тенденцию при разработке таких схем (линейность прямого хода), но не характеризует форму кривой в целом. Поэтому в дальнейшем примем первый из перечисленных терминов: генераторы пилообразного напряжения (ГПН).

Из рис. 21-1 видно, что пилообразное напряжение может быть положительным или отрицательным, а также *нарастающим* или *спадающим*. Последние два варианта различаются соотношением модулей начального и конечного напряжений прямого хода: неравенство $|U_2| > |U_1|$ характеризует нарастающее пилообразное напряжение (рис. 21-1, а, б), а неравенство $|U_2| < |U_1|$ — спадающее пилообразное напряжение (рис. 21-1, в, г).

Вообще говоря, пилообразные импульсы можно получить с помощью любого релаксатора: мультивибратора, одновибратора или блокинг-генератора. Для этого достаточно снимать выходное напряжение с времязадающего конденсатора (см. рис. 18-5, 19-2 и 20-4). Однако такой способ встречает два затруднения. Во-первых, времязадающая емкость, как правило, не имеет заземленной точки (исключения составляют схема на рис. 18-7 и ее аналоги). Во-вторых, и это главное, простейшие релаксаторы не обеспечивают необходимой линейности прямого хода. Поэтому ГПН составляют особый класс импульсных устройств и заслуживают специального рассмотрения.

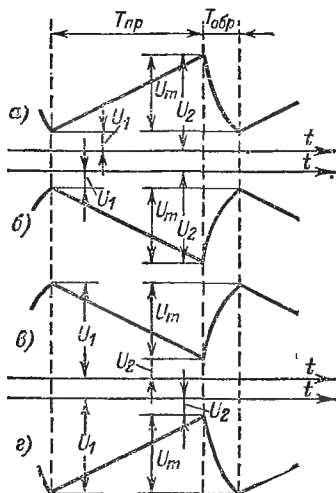


Рис. 21-1. Формы пилообразного напряжения.

Параметры генераторов. Общий анализ проведем применительно к генераторам положительного нарастающего напряжения.

В основе всех реальных ГПН лежит заряд конденсатора постоянным током. Требование постоянства зарядного тока очевидно из выражения

$$\frac{du_c}{dt} = u'_c = \frac{i_c}{C}, \quad (21-1)$$

если учесть, что идеальная линейность функции $u_c(t)$ означает условие $u'_c = \text{const}$. С этой точки зрения удобно оценивать линейность напряжения u_c либо относительной разностью скоростей заряда в начале и конце прямого хода

$$\varepsilon = \frac{U'_c(0) - U'_c(T_{\text{пр}})}{U'_c(0)}, \quad (21-2a)$$

либо (если $C = \text{const}$) относительной разностью зарядных токов:

$$\varepsilon = \frac{I_c(0) - I_c(T_{\text{пр}})}{I_c(0)}. \quad (21-26)$$

Величину ε называют *коэффициентом нелинейности* пилообразного напряжения.

Поскольку в реальных ГПН коэффициент нелинейности не превышает 5—10%, можно считать скорость заряда практически постоянной:

$$U'_c \approx \frac{U_m}{T_{\text{пр}}}, \quad (21-3a)$$

где

$$U_m = |U_2 - U_1|$$

— амплитуда пилообразного напряжения.

При расчете ГПН задаются двумя из трех величин, входящих в соотношение (21-3a). Необходимые значения зарядного тока и емкости рабочего конденсатора определяют из соотношения

$$U'_c \approx I_c/C. \quad (21-36)$$

Величину

$$\xi = \frac{U_m}{E}, \quad (21-4)$$

где E — напряжение питания, называют *коэффициентом использования напряжения питания*.

Отношение ξ/ε , которое характеризует совокупность обоих главных параметров, можно назвать *добротностью* ГПН:

$$\delta = \frac{\xi}{\varepsilon}. \quad (21-5)$$

Очевидно, что качество ГПН в целом тем выше, чем больше его добротность.

Скелетная схема и режимы работы. Методы построения линейных ГПН удобно рассмотреть с помощью общей скелетной схемы, показанной на рис. 21-2. Здесь E — источник питания; $З$ — зарядный элемент, обеспечивающий заряд конденсатора C в интервале $T_{пр}$; P — разрядный элемент, обеспечивающий заряд конденсатора C в интервале $T_{обр}$ и запертый в интервале $T_{пр}$; H — нагрузка (или чаще некоторый буферный элемент) с сопротивлением R_H .

В идеальном случае, когда отсутствуют ток нагрузки I_H и ток в запертом разрядном элементе I_P , имеют место соотношения $I_{спр} = I_{пр}$ и $I_{собр} = I_{обр} - I_{пр}$ (см. сплошные стрелки на рис. 21-2)*. В реальных схемах токи I_H и I_P неизбежны (см. пунктирные стрелки) и потому

$$I_{спр} = I_{пр} - (I_H + I_P); \quad (21-6a)$$

$$I_{собр} = (I_{обр} - I_{пр}) + I_H. \quad (21-6б)$$

Как правило, и разрядный элемент, и нагрузка представляют собой транзисторы или транзисторные схемы, токи которых можно представить в первом приближении как сумму постоянной составляющей и составляющей, пропорциональной напряжению [см. (2-50)]:

$$i_H = I_{H0} + u_c/R_H; \quad (21-7a)$$

$$i_P = I_{P0} + u_c/R_P. \quad (21-7б)$$

В частных случаях одной из этих составляющих можно пренебрегать (например, в случае кремниевых транзисторов — токами I_{H0} или I_{P0}). Сумму обоих токов запишем в виде¹

$$i_{ш} = i_H + i_P = I_{ш0} + u_c/R_{ш}, \quad (21-7в)$$

где $I_{ш0} = I_{H0} + I_{P0}$; $R_{ш} = R_H \parallel R_P$. Индекс «ш» соответствует шунтирующей роли элементов P и H по отношению к C .

Из выражения (21-6б) очевидно, что обратный ход имеет место только при условии $I_{обр} > I_{пр}$ и что он будет короче прямого хода, если $I_{обр} > 2I_{пр}$ (токами I_H и I_P пренебрегаем). Положим, что между временами $T_{пр}$, $T_{обр}$ и токами $I_{спр}$, $I_{собр}$ имеет место обратная пропорциональность:

$$T_{пр} : T_{обр} = (I_{собр})_{ср} : I_{спр},$$

* Для токов зарядного и разрядного элементов приняты обозначения $I_{пр}$ и $I_{обр}$, чтобы их можно было использовать и в генераторах спадающего напряжения, в которых прямому ходу соответствует разряд, а обратному ходу — заряд конденсатора (см. рис. 21-1, в и г).

¹ Строго говоря, при $u_c = 0$ ток $i_{ш} = 0$. Однако в дальнейшем это обстоятельство для простоты не учитывается.

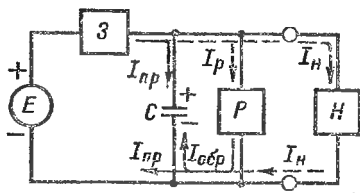


Рис. 21-2. Общая скелетная схема генератора нарастающего положительного напряжения.

где $(I_{\text{обр. ср}})$ — среднее значение тока (поскольку этот ток, как правило, меняется во время разряда). В большинстве случаев желательно иметь обратный ход как можно короче. Поэтому, исходя из приведенной пропорции и учитывая (21-6), можно записать неравенство

$$I_{\text{обр. ср}} \geq I_{\text{пр}} \left(1 + \frac{T_{\text{пр}}}{T_{\text{обр}}} \right), \quad (21-8)$$

где $T_{\text{обр}}$ — максимально приемлемое значение. Обычно $T_{\text{пр}} \gg T_{\text{обр}}$, и тогда $I_{\text{обр. ср}} \gg I_{\text{пр}}$.

Разрядные элементы, а вместе с ними и ГПН в целом могут работать в двух существенно разных режимах: *режиме самовозбуждения* и *режиме с посторонним возбуждением*. Режим самовозбуждения характерен тем, что разрядный элемент представляет собой пороговое устройство, которое срабатывает при некотором напряжении U_2 и разряжает конденсатор до нулевого напряжения (или в общем случае до напряжения $U_1 < U_2$), после чего снова запирается на время прямого хода. Таким образом, в этом режиме амплитуда пилообразного напряжения оказывается заданной величиной, а время прямого хода согласно (21-3а) — функцией скорости U'_c .

Режим с посторонним возбуждением характерен тем, что разрядный элемент представляет собой ключ, управляемый некоторым импульсным устройством (мультивибратор, триггер, одновибратор). В этом случае заданным оказывается интервал $T_{\text{пр}}$, а амплитуда согласно (21-3а) — функцией скорости U'_c .

Методу постороннего возбуждения обычно свойственно большее выходное напряжение, что весьма важно для ГПН¹. Кроме того, этот метод более гибок; в частности, наряду с непрерывными колебаниями он легко обеспечивает генерацию однократных импульсов, т. е. *жадущий режим* (для этого достаточно, чтобы разрядный элемент был нормально открыт и запирался прямоугольным сигналом на время прямого хода).

Разновидности генераторов. ГПН принято классифицировать по тем методам, с помощью которых достигается линейность прямого хода. Если обозначить через R сопротивление зарядного элемента постоянному току, то из рис. 21-2 следует:

$$i_{\text{пр}} = \frac{E - u_c}{R}.$$

Если далее принять для суммарного шунтирующего тока выражение (21-7в) и подставить значения $i_{\text{пр}}$ и $i_{\text{ш}}$ в формулу (21-6а), то после преобразований получим:

$$i_c = \left(\frac{E}{R} - I_{\text{ш0}} \right) - \frac{u_c}{R \parallel R_{\text{ш}}}. \quad (21-9)$$

¹ Генерация малых выходных напряжений с последующим усилением неприемлема, поскольку любой усилительный каскад вносит такие нелинейные искажения, которые, как правило, превышают заданную величину ϵ .

Тогда согласно (21-1)

$$u_c' = \frac{1}{C} \left[\left(\frac{E}{R} - I_{ш0} \right) - \frac{u_c}{R \parallel R_{ш}} \right]. \quad (21-10)$$

Выражение (21-10) есть дифференциальное уравнение прямого хода. Не решая этого уравнения, рассмотрим пути, которые способны обеспечить возможно большее постоянство его правой части и тем самым высокую линейность функции $u_c(t)$.

1. Наиболее простой и очевидный путь (рис. 21-3, а) — использование постоянного зарядного сопротивления R и достаточно большого напряжения $E \gg U_2$, где U_2 — максимальное напряжение на конденсаторе (см. рис. 21-1). Поскольку цепочка RC в данном случае интегрирующая [62, гл. 10], назовем

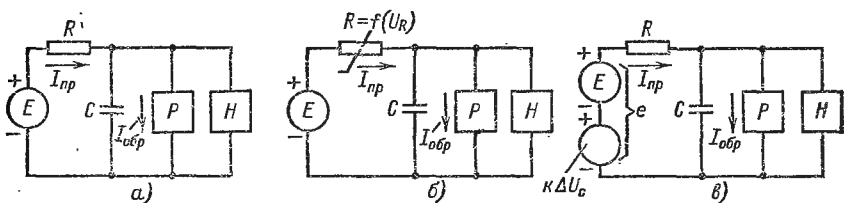


Рис. 21-3. Скелетные схемы генераторов.

а — простейшего с интегрирующей цепочкой; б — с параметрическим стабилизатором тока; в — со следящей связью

генераторы, построенные на таком принципе, *простейшими ГПН с интегрирующей цепочкой*.

Введем обозначения:

$\vartheta_{шI} = \frac{I_{ш0}}{E/R}$ — фактор шунтирующего тока;

$\vartheta_{шR} = R/R_{ш}$ — фактор шунтирующего сопротивления;

$\vartheta_{ш1} = U_1/E$ — фактор начального напряжения (значение U_1 — см. на рис. 21-1).

Затем, определяя из (21-9) зарядные токи в моменты $t = 0$ (когда $u_c = U_1$) и $t = T_{пр}$ (когда $u_c = U_2$), подставляя эти токи в определение (21-26) и учитывая соотношение $U_2 - U_1 = U_m$, получаем коэффициент нелинейности в виде

$$\epsilon = \frac{U_m}{E} \vartheta, \quad (21-11a)$$

где

$$\vartheta = \frac{1 + \vartheta_{шR}}{1 - \vartheta_{шI} - \vartheta_{ш1}(1 + \vartheta_{шR})}.$$

Величина ϑ является поправочным коэффициентом, характеризующим неидеальность ГПН, т. е. влияние на коэффициент нелинейности таких «паразитных» факторов, как начальное напряжение U_1 и разного рода шунты. В идеальном случае, когда $\vartheta_{ш1}$,

$$\vartheta_{\text{ш}R}, \vartheta_{\text{ш}I} = 0 \text{ и } \vartheta = 1,$$

$$\epsilon_0 = \frac{U_m}{E}. \quad (21-11б)$$

Как видим, при значительных величинах U_m и малых ϵ необходимое напряжение E может оказаться весьма большим. Например, если $\vartheta = 1$; $U_m = 10$ В и $\epsilon = 1\%$ (нередко требуются гораздо меньшие значения ϵ), то $E = 1000$ В, что совершенно неприемлемо.

Что касается остальных параметров, то коэффициент использования питающего напряжения выражается общей формулой (21-4), а добротность согласно (21-5) равна:

$$\delta = \frac{1 - \vartheta_{\text{ш}I}}{1 + \vartheta_{\text{ш}R}} - \vartheta_{\text{ш}I}, \quad (21-12)$$

т. е. она всегда меньше единицы.

2 Второй путь (рис. 21-3, б) состоит в использовании нелинейного зарядного элемента, у которого сопротивление постоянному току R уменьшается с уменьшением напряжения так, что ток остается почти неизменным. В сущности речь идет о том, чтобы зарядный элемент имел достаточно большое дифференциальное сопротивление, т. е. был стабилизатором тока. Генераторы, построенные по такому принципу, назовем *ГПН с параметрическим стабилизатором тока*.

Введем обозначение R_I для дифференциального сопротивления зарядного элемента и положим $R_I = \text{const}$ в рабочем диапазоне напряжений (от $E - U_1$ до $E - U_2$). Тогда приращение зарядного тока конденсатора во время прямого хода можно записать так:

$$\Delta I_c = I_c(0) - I_c(T_{\text{пр}}) = \Delta I_{\text{пр}} + \Delta I_{\text{ш}} = \frac{U_m}{R_I} + \frac{U_m}{R_{\text{ш}}}.$$

Далее обозначим через

$$I_{\text{пр}}(0) = \frac{E}{R(0)}$$

начальный (максимальный) ток параметрического стабилизатора, т. е. ток при $U_1 = 0$, когда все напряжение E падает на стабилизаторе; практически этот ток мало меняется в течение рабочего цикла. Подставляя $I_{\text{пр}}(0)$ в выражение (21-9), получаем начальный ток конденсатора:

$$I_c(0) = I_{\text{пр}}(0) - I_{\text{ш}0} - \frac{U_1}{R_I \parallel R_{\text{ш}}}.$$

Используя значения ΔI_c и $I_c(0)$, можно привести коэффициент нелинейности (21-2б) к виду

$$\epsilon = \frac{U_m}{E_I} \vartheta, \quad (21-13а)$$

где

$$\vartheta = \frac{1 + \vartheta_{\text{ш}R}}{1 - \vartheta_{\text{ш}I} - \vartheta_{\text{ш}I}(1 + \vartheta_{\text{ш}R})}.$$

В идеальном случае, когда $\vartheta_{ш1}, \vartheta_{шR}, \vartheta_{шR}' = 0$ и $\vartheta = 1$,

$$\epsilon_0 = \frac{U_m}{E_i}. \quad (21-136)$$

Эквивалентная э. д. с. E_i представляет собой произведение прямого тока на внутреннее сопротивление параметрического стабилизатора:

$$E_i = I_{пр}(0) R_i. \quad (21-14)$$

Факторы $\vartheta_{ш1}, \vartheta_{шR}, \vartheta_{ш1}'$ имеют такую же структуру, как и в простейших ГПН, если заменить E на E_i и R на R_i :

$$\vartheta_{ш1} = \frac{I_{ш0}}{E_i/R_i}; \quad \vartheta_{шR} = \frac{R_i}{R_{ш}}; \quad \vartheta_{ш1}' = \frac{U_1}{E_i}.$$

Как видим, для заданных величин U_m и ϵ нужно обеспечить достаточно большое значение эквивалентной э. д. с. E_i . Обычно это не представляет труда, так как сопротивление R_i у стабилизаторов тока велико и необходимое значение E_i получается при сравнительно небольших токах и напряжениях питания¹.

Например, если по-прежнему $U_m = 10$ В; $\epsilon = 1\%$ и $\vartheta = 1$, то $E_i = 1000$ В. Согласно (21-14) для реализации такой величины при $R_i = 0,5$ МОм требуется $I_{пр}(0) = 2$ мА.

Разделив (21-4) на (21-13а) и учтя (21-14), получим добротность данного варианта ГПН:

$$\delta = \frac{E_i}{E} \left[\frac{1 - \vartheta_{ш1}}{1 + \vartheta_{шR}} - \vartheta_{ш1}' \right]; \quad (21-15)$$

она, как правило, много больше единицы. Так, если в условиях последнего примера положить $E = 20$ В, то $\delta = 50$.

Поскольку в любом стабилизаторе тока $R_i \gg R(0)$, то фактор шунтирующего сопротивления $\vartheta_{шR}$ оказывается в данном типе ГПН значительно больше, чем в простейших ГПН, при том же зарядном токе. Соответственно возрастает влияние $R_{ш}$ на коэффициент нелинейности. Для того чтобы уменьшить это влияние, приходится принимать меры к увеличению $R_{ш}$, руководствуясь желательным соотношением $R_{ш} \gg R_i$.

Из формулы (21-13а), полагая $R_i \rightarrow \infty$, нетрудно получить:

$$\epsilon = \frac{U_m}{I_{пр} R_{ш}} \left(1 - \vartheta_{ш1} - \frac{U_1}{I_{пр} R_{ш}} \right)^{-1}.$$

Значит, даже в случае идеального стабилизатора тока коэффициент нелинейности ГПН не равен нулю, а определяется сопротивлением шунтов, что, конечно, нежелательно из соображений неустойчивости.

3. Третий путь (рис. 21-3, в) состоит в замене постоянного напряжения питания E меняющимся напряжением вида

$$e = E + k(u_c - U_1), \quad (21-16)$$

¹ Разумеется, должно выполняться условие $E > U_2$, с тем чтобы зарядный элемент находился в рабочем режиме даже при максимальном напряжении на конденсаторе.

где $k = de/du_c$ — коэффициент пропорциональности. Тогда из (21-9), подставляя e вместо E и полагая сначала $u_c = U_1$, а затем $u_c = U_2$, получаем:

$$I_c(0) = \frac{E}{R} - I_{\text{ш0}} - \frac{U_1}{R \parallel R_{\text{ш}}};$$

$$I_c(T_{\text{пр}}) = \frac{E + kU_m}{R} - I_{\text{ш0}} - \frac{U_2}{R \parallel R_{\text{ш}}};$$

соответственно из определения (21-26) следует:

$$\varepsilon = \frac{U_m(1-k)}{E} \vartheta, \quad (21-17a)$$

где

$$\vartheta = \frac{1 + \vartheta_{\text{ш}R}/(1-k)}{1 - \vartheta_{\text{ш}I} - \vartheta_{\text{ш}1}(1 + \vartheta_{\text{ш}R})}.$$

В идеальном случае, когда $\vartheta_{\text{ш}1}$, $\vartheta_{\text{ш}R}$, $\vartheta_{\text{ш}I} = 0$ и $\vartheta = 1$,

$$\varepsilon_0 = \frac{U_m}{E} (1-k). \quad (21-176)$$

В формуле (21-17a) факторы $\vartheta_{\text{ш}1}$, $\vartheta_{\text{ш}R}$, $\vartheta_{\text{ш}I}$ имеют те же значения, что и в простейшей схеме (с. 591). Однако в данном случае в выражении для ϑ фактор $\vartheta_{\text{ш}R}$ делится на малую величину $1-k$, что эквивалентно резкому увеличению сопротивления R ; это заставляет соответственно увеличивать сопротивление $R_{\text{ш}}$ (см. аналогичные замечания в предыдущем пункте).

Сравнивая выражения (21-176) и (21-116), легко заметить, что в данном случае при прочих равных условиях необходимое напряжение E в $1/(1-k)$ раз меньше, чем в простейшей схеме. Например, если $k = 0,98$, то вместо напряжения $E = 1000$ В, необходимого в простейшей схеме, потребуется всего лишь $E = 20$ В. Аналогичный выигрыш можно обеспечить и при $\vartheta \neq 1$, выбрав соответствующее значение k .

Зависимость (21-16) предполагает, что напряжение e «следит» за изменениями напряжения u_c с помощью специального следящего устройства. Поэтому такой метод стабилизации тока назовем *следящей связью* (английское boot-strap), а соответствующие генераторы — ГПН со следящей связью¹.

Деля (21-4) на (21-17a), получаем добротность ГПН в виде

$$\delta = \frac{1 - \vartheta_{\text{ш}I} - \vartheta_{\text{ш}1}(1 + \vartheta_{\text{ш}R})}{1 - k + \vartheta_{\text{ш}R}}. \quad (21-18)$$

Как видим, в зависимости от величины k добротность может быть сколь угодно большой; практически она составляет до 100 и более.

¹ В частном случае следящей связи, когда $E = 0$, т. е. когда «стабилизируется» (поддерживается) нулевой начальный ток, получается обычная обратная связь по напряжению сигнала (см. рис. 9-8 и 9-9).

Анализ выражений (21-11) — (21-18) приводит к следующим выводам. Простейшие ГПН пригодны при сравнительно малых выходных напряжениях (десятые доли вольта) и сравнительно больших коэффициентах нелинейности (несколько процентов). ГПН с нелинейным зарядным элементом при наличии подходящего стабилизатора тока могут обеспечить гораздо лучшие параметры (выходные напряжения в несколько вольт и нелинейность до 1%); однако шунтирующее сопротивление $R_{ш}$ ставит принципиальный предел повышению линейности. ГПН со следящей связью наиболее совершенны и универсальны; они могут обеспечить нелинейности порядка десятых долей процента, а выходные напряжения ограничены лишь допустимыми параметрами транзисторов. Дополнительные особенности рассмотренных вариантов будут отмечены в последующих параграфах.

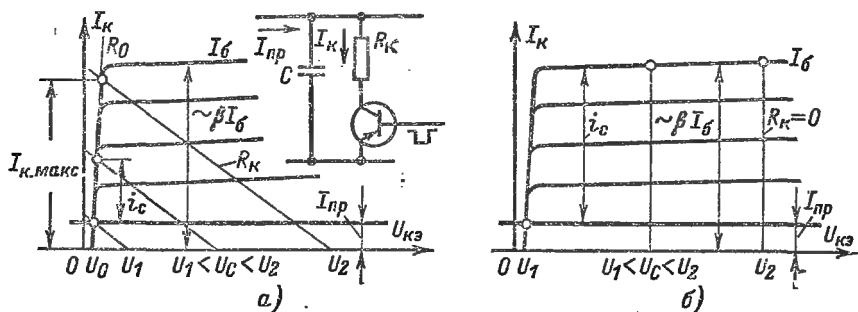


Рис. 21-4. Работа разрядного элемента в насыщенном (а) и ненасыщенном (б) режимах.

Обратный ход. Как уже отмечалось, в ГПН с самовозбуждением роль разрядных элементов играют пороговые устройства. В зависимости от конкретной схемы такого устройства длительность обратного хода и характер разряда могут быть разными. Что касается ГПН с посторонним возбуждением, то у них разрядные элементы реализуются с помощью насыщенных или ненасыщенных транзисторных ключей, которым свойственны вполне определенные закономерности. Поэтому имеет смысл рассмотреть обратный ход на примере ГПН с посторонним возбуждением, тем более что этот режим более распространен на практике.

Обычно разрядник представляет собой совокупность транзистора, включенного по схеме ОЭ, и резистора R_K , обеспечивающего насыщенный режим транзистора в течение всего разряда (рис. 21-4, а). Возможен другой вариант разрядника, в котором резистор R_K отсутствует, а следовательно, транзистор работает в активном режиме, поскольку на нем падает все напряжение U_C .

Независимо от наличия резистора запирающий разрядник можно представить генератором тока I_{K0} , зашунтированным сопротивлением запирающего R_3 (см. § 15-3). Ток I_{K0} входит составной частью в ток $I_{ш}$, а сопротивление R_3 — составной частью в $R_{ш}$. Если используется инверсное включение транзистора, ток I_{K0} следует заменить меньшей величиной $(\beta_I/\beta_N) I_{K0}$ (см. § 15-2).

В открытом состоянии свойства разрядника зависят от режима транзистора, т. е. от наличия или отсутствия резистора R_k .

В случае насыщенного ключа (рис. 21-4, а) всегда выполняются условия $U_0 \ll U_1$ и $R_0 \ll R_k$, где U_0 — остаточное напряжение на насыщенном транзисторе и R_0 — сопротивление участка коллектор — эмиттер в режиме насыщения (сопротивление отпирания, см. § 15-3). Поэтому можно считать, что $U_{кэ} = 0$ и что конденсатор C разряжается через R_k с постоянной времени, близкой к CR_k .

Время разряда, учитывая экспоненциальный характер процесса, составляет:

$$T_{обр} \approx 2,3CR_k. \quad (21-19)$$

За это время напряжение U_c меняется от максимального значения U_2 до минимального

$$U_1 = I_{пр}R_k, \quad (21-20)$$

которое соответствует условию $I_c = 0$.

Будем считать средний ток разряда равным $0,5 I_{к. макс.}$. Тогда из неравенства (21-8) получаем:

$$I_{к. макс.} \geq 2I_{пр} \left(1 + \frac{T_{пр}}{T_{обр}} \right). \quad (21-21)$$

Выбрав ток $I_{к. макс.}$, можно оценить необходимое значение R_k из соотношения $I_{к. макс.} = U_2/R_k$, а также ток базы из условия насыщения $\beta I_b > I_{к. макс.}$.

В случае ненасыщенного ключа (рис. 21-4, б) конденсатор разряжается почти постоянным током

$$I_k - I_{пр} \approx \beta I_b - I_{пр}.$$

Время разряда можно принять равным отношению начального заряда на конденсаторе (CU_2) к разряжающему току:

$$T_{обр} \approx \frac{CU_2}{I_k - I_{пр}} \approx \frac{CU_m}{\beta I_b} \quad (21-22)$$

(второе приближение действительно, если $\beta I_b \gg I_{пр}$).

В самом конце разряда транзистор насыщается, так как ток коллектора (при неизменном токе базы) уменьшается до значения $I_{пр}$. При этом минимальное напряжение U_1 , близкое к U_0 , будет пренебрежимо мало.

Как известно, допустимый ток в активном режиме в несколько раз меньше, чем в режиме насыщения: $(I_{к. доп})_{акт} \ll (I_{к. доп})_{нас}$ (см. § 15-8). Соответственно время разряда при активном режиме оказывается заметным, чем при насыщении. Это нетрудно показать, исходя из выражений (21-22) и (21-19), если в первом положить $(I_{к. доп})_{акт} \geq I_{пр}$, а во втором $R_k \approx U_2/(I_{к. доп})_{нас}$. Поэтому в качестве разрядных элементов используются, как правило, насыщенные ключи, хотя напряжение U_1 у них больше и иногда составляет 10—20% U_2 .

Особенности генераторов спадающего напряжения. Достаточно разрядить емкость стабильным током и быстро зарядить ее, чтобы генератор нарастающего напряжения превратился в генератор спадающего напряжения (ср. кривые на рис. 21-1, а и б). Скелетная схема последнего не отличается от схемы на рис. 21-2, но функции элементов Z и P меняются местами: теперь зарядный элемент является либо пороговым устройством (режим самовозбуждения), либо ключом (режим постороннего возбуждения), а разрядный элемент — источником стабильного тока, выполненным по одному из трех вариантов, рассмотренных выше. Соответственно элемент P осуществляет прямой ход, а элемент Z — обратный.

Специфика спадающего напряжения состоит в следующем.

1. Начальный ток разрядного элемента (в отличие от начального тока зарядного элемента на рис. 21-2) определяется не разностью напряжений E и u_c , а напряжением u_c :

$$i_{np} = \frac{u_c}{R}.$$

2. Ток нагрузки при протекании через емкость не вычитается, а складывается с током прямого хода, т. е. в выражениях (21-6) следует изменить знак перед током I_n .

3. Ток через элемент 3 во время прямого хода описывается выражением

$$i_3 = I_{30} + \frac{E - u_c}{R_3},$$

в отличие от выражения (21-7б) для соответствующего тока через элемент P .

С учетом отмеченной специфики получаем для прямого хода вместо формулы (21-9) формулу

$$i_c = \frac{u_c}{R \parallel R_{ш}} - \Delta I_{ш} - \frac{E}{R_3}, \quad (21-23)$$

где $\Delta I_{ш} = I_{30} - I_{н0}$ — разность постоянных составляющих шунтирующих токов, а параметр R может быть либо постоянным, либо переменным в зависимости от типа ГПН.

Подставляя в (21-23) значения $U_c(0) = U_1$ и $U_c(T_{пр}) = U_2$, нетрудно получить из определения (21-2б) коэффициенты нелинейности для трех вариантов ГПН. Эти коэффициенты выражаются теми же формулами (21-11), (21-13) и (21-17), что и в случае генераторов нарастающего напряжения. Однако поправочные коэффициенты ϑ имеют другие значения. А именно, для простейших ГПН

$$\vartheta = \frac{1 + \vartheta_{шR}}{\vartheta_{ш1}(1 + \vartheta_{шR}) - \Delta\vartheta_{ш1} - \vartheta_{ш3}}; \quad (21-24a)$$

для ГПН с параметрическим стабилизатором тока

$$\vartheta = \frac{1 + \vartheta_{шR}}{\vartheta_{ш1}\vartheta_{шR} - \Delta\vartheta_{ш1} - \vartheta_{ш3}}; \quad (21-24б)$$

для ГПН со следящей связью

$$\vartheta = \frac{1 + \vartheta_{шR}/(1-k)}{\vartheta_{ш1}(1 + \vartheta_{шR}) - \Delta\vartheta_{ш1} - \vartheta_{ш3}}. \quad (21-24в)$$

Частные факторы $\vartheta_{ш1}$, $\vartheta_{шR}$ и $\vartheta_{ш3}$ имеют ту же структуру, что в генераторах нарастающего напряжения:

$$\Delta\vartheta_{ш1} = \frac{\Delta I_{ш}}{E/R}; \quad \vartheta_{шR} = \frac{R}{R_{ш}}; \quad \vartheta_{ш1} = \frac{U_1}{E} \approx 1,$$

где в случае ГПН с параметрическим стабилизатором тока $E = E_1 = I_{пр} R_1$ и $R = R_1$. Фактор $\Phi_{шв}$ является частным случаем фактора $\Phi_{шR}$, если заменить R на R_3 .

Таким образом, в идеальном случае ($\Phi = 1$) коэффициенты нелинейности для нарастающего и спадающего напряжений одинаковы. В реальных случаях у генераторов спадающего напряжения величины Φ при прочих равных условиях несколько больше.

Ниже рассматриваются некоторые типичные примеры ГПН. Более подробно схемы ГПН и их расчет рассмотрены в работах [164, 165, 168, 169, 171].

21-3. ПРОСТЕЙШИЕ ГЕНЕРАТОРЫ С ИНТЕГРИРУЮЩЕЙ ЦЕПОЧКОЙ

Типичным примером простейшего ГПН является схема на рис. 21-5, которая отличается от аналогичной схемы мультивибратора (см. рис. 18-7) тем, что выходное напряжение снимается с конденсатора C , а триггер с эмиттерной связью играет роль разрядного элемента¹. Очевидно, что данный генератор относится к ГПН с самовозбуждением. Чтобы исключить составляющую напряжения U_1 (рис. 18-8), используются переходный конденсатор C_1 и фиксирующий диод D . Эти схемные элементы обеспечивают нулевой начальный уровень пилообразного напряжения на выходе и не оказывают существенного влияния на форму сигнала; поэтому при анализе основных процессов ими можно пренебречь. В остальном структура данного ГПН соответствует скелетной схеме на рис. 21-2.

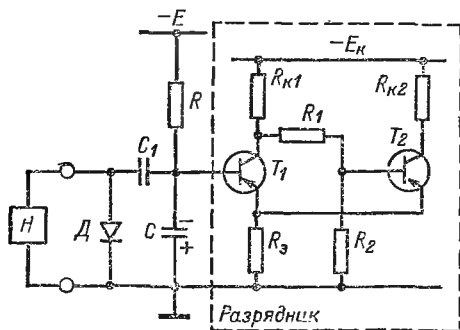


Рис. 21-5. Простейший генератор с интегрирующей цепочкой и разрядным триггером.

Напряжение u_c меняется в пределах от $U_1 = U_{III}$ до $U_2 = U_{II}$ (рис. 18-8), т. е.

$$U_m = U_{II} - U_{III}.$$

Если в формулах (17-5) и (17-8) положить $g_2' \ll \gamma_1 g_3$; $\beta_2 \gamma_1 \gg 1$ (что обычно соблюдается) и согласно (17-21а) принять $g_{k1} = g_{k2} = g_k$, то

$$U_2 = U_{II} \approx \frac{E_k}{1 + g_3/2g_k};$$

$$U_1 = U_{III} \approx \frac{E_k}{1 + g_3/g_k},$$

¹ Если триггер используется в виде интегральной схемы, то весь разрядный элемент представляет собой одну деталь.

откуда следует:

$$U_m = E_k \frac{g_3/g_k}{(1 + g_3/g_k)(2 + g_3/g_k)}. \quad (21-25)$$

Дифференцируя функцию U_m по аргументу g_3/g_k , легко убедиться, что она имеет максимум:

$$U_{m \text{ макс}} \approx 0,17 E_k \quad (21-26)$$

при $g_3/g_k = \sqrt{2}$. В этом оптимальном случае $U_1 = U_{III} \approx 0,4 E_k$. Для распространенного диапазона $E_k = 5 \div 15$ В имеем $U_{m \text{ макс}} = 0,8 \div 2,5$ В; $U_1 = 2 \div 6$ В. Максимум функции $U_m(g_3/g_k)$ выражен неярко (рис. 21-6), поэтому можно уменьшать или увеличивать отношение g_3/g_k примерно в 3 раза по сравнению с оптимальным значением, сохраняя амплитуду U_m в пределах 70% максимальной.

Если пренебречь шутирующим влиянием нагрузки и разрядного элемента, т. е. считать $\phi_{шI} = \phi_{шR} = 0$, то из выражения (21-11a) нетрудно получить коэффициент нелинейности в виде

$$\varepsilon = \frac{g_3/g_k}{(2 + g_3/g_k) \left[\frac{E}{E_k} (1 + g_3/g_k) - 1 \right]}. \quad (21-27)$$

Эта функция также имеет слабо выраженный максимум при значениях $E > E_k$ (рис. 21-6), в чем легко убедиться путем дифференцирования. Максимальное значение ε получается в области малых значений g_3/g_k , не превышающих $\sqrt{2}$. В области $g_3/g_k > \sqrt{2}$ коэффициент нелинейности монотонно уменьшается с ростом g_3/g_k , а также с увеличением отношения E/E_k .

С помощью выражений (21-25) и (21-27) получаем добротность ГПН:

$$\delta = 1 - \frac{1}{(E/E_k)(1 + g_3/g_k)}; \quad (21-28)$$

как видим, она возрастает с увеличением E/E_k и g_3/g_k (рис. 20-6).

Анализируя кривые на рис. 21-6, приходим к следующим выводам:

1. С точки зрения линейности нужно делать отношение E/E_k как можно большим, а с точки зрения добротности достаточно обеспечить $E/E_k = 2 \div 5$. Значения $E \leq E_k$ неприемлемы, так как им соответствуют параметры $\varepsilon > 15\%$ и $\delta < 0,8$. Ограничения на значение E/E_k накладываются, с одной стороны, значением выходного напряжения¹, а с другой — разумным значением питающего напряжения E .

2. Отношение g_3/g_k с точки зрения линейности и добротности желательно делать большим. Однако значения $g_3/g_k > 3$ нецелесо-

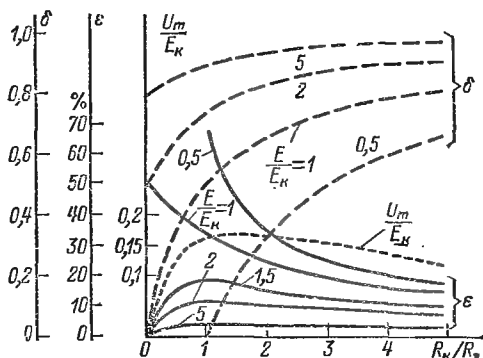


Рис. 21-6. Зависимость амплитуды, коэффициента нелинейности и добротности генератора (см. рис. 21-5) от соотношения напряжений E/E_k и проводимостей g_3/g_k .

¹ При заданной амплитуде U_m напряжение E_k должно быть достаточно большим в силу зависимости (21-25).

собразны, поскольку при этом выигрывают в параметрах ϵ и δ невелик, но происходит заметное уменьшение выходного напряжения, а также могут нарушиться условия триггерного режима (17-15) и (17-21).

В диапазонах $E/E_k = 2 \div 5$ и $g_g/g_k = 2 \div 3$ получаем типичные для данного ГПН параметры: $\epsilon = 3 \div 10\%$ и $\delta = 0,8 \div 0,95$. Шунтирующее влияние нагрузки и входного сопротивления запертого транзистора T_1 ухудшает параметры ϵ и δ .

Помимо триггера с эмиттерной связью в качестве разрядного элемента могут использоваться другие пороговые схемы (например, в [168] описаны ГПН с разряжающим мультивибратором).

21-4. ГЕНЕРАТОРЫ С ПАРАМЕТРИЧЕСКИМ СТАБИЛИЗАТОРОМ ТОКА

В данном варианте ГПН, как отмечалось в § 21-2, главным требованием к зарядному элементу является возможно большее дифференциальное сопротивление. Наиболее простым элементом такого типа может служить транзистор, включенный по схеме ОБ с заданным током эмиттера (рис. 4-10, а). В таком режиме дифференциальное сопротивление $R_i = r_k$, а эквивалентная э. д. с. $E_i = I_{пр} R_i = I_k r_k$. Учитывая зависимость (4-24), приходим к выводу, что э. д. с. E_i не зависит от тока, а является функцией только физических параметров базы и коллекторного напряжения (для ступенчатого перехода $E_i \sim \sqrt{U_k}$)*.

Справочные данные показывают, что маломощным низкочастотным транзисторам свойственны (при $U_k = 5$ В) значения $I_k r_k = 300 \div 500$ В, а маломощным высокочастотным — от 1000 до 2000 В и более¹. Таким образом, ГПН данного типа могут обеспечить коэффициенты нелинейности 1—5% даже при таких сравнительно больших амплитудах, как $U_m = 10 \div 20$ В (в этих оценках не учитываются шунтирующие токи). Что касается максимального значения U_m , то оно определяется напряжением $U_{кб. доп}$ и может составить 40—60 В и больше.

Сопротивление r_k играет важную роль не только как составляющая э. д. с. E_i , но также и как составляющая фактора $\Phi_{шR}$ в выражении (21-13а)**. При токе 1 мА значение r_k обычно лежит в пределах 0,5—2 МОм.

Такое же и даже меньшее значение часто имеет сопротивление $R_{ш}$; тогда $\Phi_{шR} \geq 1$ и может существенно влиять на коэффициент нелинейности. Напомним также, что r_k зависит от температуры (рис. 4-20), поэтому коэффициент ϵ оказывается нестабильным.

* Последнее обстоятельство означает, что при расчете коэффициента нелинейности нужно использовать величину $E_i = I_k r_k$, соответствующую минимальному напряжению U_k .

¹ Последние цифры характерны также для некоторых низкочастотных транзисторов средней мощности, например, 1Т403 с допустимым напряжением до 100 В и токами до 1 А.

** Фактор $\Phi_{шI}$ с учетом (21-14) не зависит от r_k .

должны быть усилителем (рис. 21-8, б)¹. Соответственно первую разновидность назовем ГПН с повторительной следящей связью, а вторую разновидность — ГПН с усилительной следящей связью. Рассмотрим отдельно особенности каждой разновидности.

Генераторы с повторительной связью. Типичный пример такого ГПН показан на рис. 21-9 вместе с временными диаграммами.

В схеме генерируется отрицательное нарастающее напряжение. Транзистор T_1 , играющий роль управляемого ключа (режим постороннего возбуждения), нормально открыт и насыщен. При этом в цепи $E_k - D - R - T_1$ протекает ток

$$I(0) \approx E_k / (R_d + R + R_k), \quad (21-32)$$

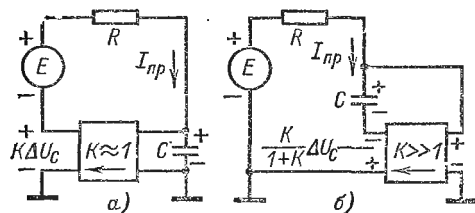


Рис. 21-8. Скелетные схемы генераторов с повторительной (а) и усилительной (б) следящей связью.

U_c . Эмиттерный повторитель, выполненный на транзисторе T_2 , почти полностью передает приращение ΔU_c через большую емкость C_0 в точку а, где оно складывается с начальным напряжением $U_a(0) = E_k - I(0) R_d$. Таким образом, работа генератора отвечает скелетной схеме на рис. 21-8, а, если положить $I_{np} = I(0)$; $E = U_a(0)$; $U_1 = I(0) R_k$ и $k = K$, где K — коэффициент передачи повторителя. Параметры ГПН в и б можно рассчитать по общим формулам (21-17) и (21-18).

Остановимся на некоторых особенностях схемы, которые следует иметь в виду при ее разработке. Роль нагрузки в данном случае играет входное сопротивление повторителя. Согласно (9-16)

$$R_{вх} \approx (1 + \beta) (r_k^* \parallel R_g \parallel R_n). \quad (21-33)$$

Следовательно, при условии $R_n \gg r_k^* \parallel R_g$ влиянием истинной нагрузки R_n на коэффициент нелинейности можно пренебречь. Следует, однако, иметь в виду, что в начале прямого хода, когда диод D еще не заперт², эквивалентное сопротивление $R_d \parallel R_n$ может быть гораздо меньше, чем R_n .

¹ Для этой разновидности, включающей в себя фантастрыоны, часто применяют термин «ГПН с емкостной обратной связью», весьма неудачный с точки зрения классификации.

² Запирание диода происходит при условии $\Delta U_a = U_d$. Учитывая соотношения $\Delta U_a = K \Delta U_c$ и $\Delta t \approx \Delta U_c / U'_c$, легко записать интервал времени, по истечении которого диод будет заперт:

$$\Delta t_d \approx \frac{U_d}{K U'_c} \approx \frac{U_d}{U_m} T_{np}.$$

В этом интервале времени входное сопротивление повторителя будет существенно зависеть от R_d , увеличиваясь по мере запираания диода. Методы повышения входного сопротивления путем использования сложных повторителей рассмотрены в § 9-4.

Величина коэффициента нелинейности зависит помимо $R_{вх}$ от начального тока базы $I_{б2}(0)$, который играет роль тока $I_{н0}$ в общей теории ГПН [см. (21-7а)]. Обычно выбирают ток базы из условия $I_{б2}(0) \gg I_{к0}$, обеспечивающего стабильный режим повторителя¹. Тогда начальный эмиттерный ток должен иметь значение

$$I_{э2}(0) = (1 + \beta) I_{б2}(0).$$

В отсутствие дополнительного источника питания $+E_э$ ток $I_{э2}(0)$ трудно, а чаще всего невозможно обеспечить, поскольку потенциал эмиттера в этом случае равен нулю, а начальный потенциал базы $U_{б2}(0)$ часто составляет доли вольта.

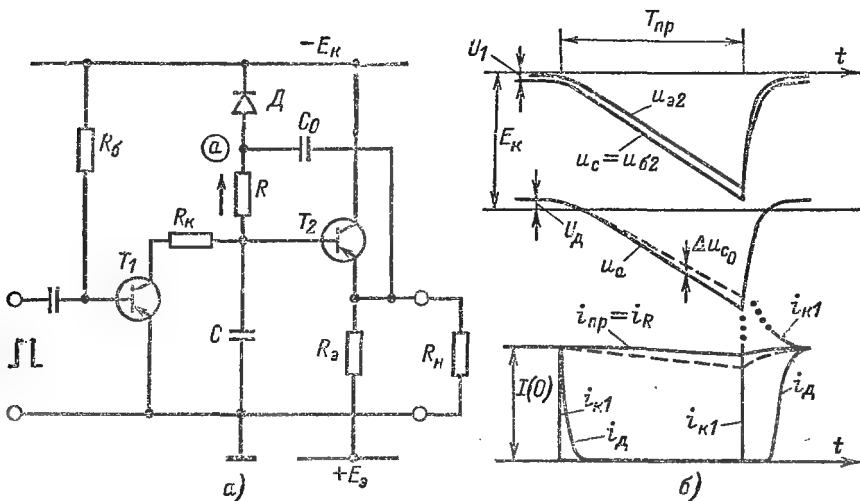


Рис. 21-9. Генератор с повторительной следящей связью.

а — принципиальная схема; б — временные диаграммы.

При наличии дополнительного источника $+E_э$ с достаточно большим напряжением (несколько вольт) ток $I_{э2}(0)$ практически равен $E_э/R_э$, так как потенциал эмиттера ненамного отличается от $U_{б2}(0)$. Следовательно, необходимый ток $I_{э2}(0)$ можно задать, используя сравнительно большое сопротивление $R_э$, что обеспечивает стабильность повторителя и его высокое входное сопротивление.

Амплитуда выходного напряжения U_m ограничена условием $U_{б2} = U_c < E_k$ с тем, чтобы коллекторный переход транзистора T_2 работал с обратным смещением. Поэтому $U_{2\max} \approx E_k$ и $U_{m\max} \approx E_k - U_1$. Если $U_1 \ll E_k$, как обычно бывает, то $U_{m\max} \approx E_k$ и коэффициент использования ξ согласно (21-4) близок к единице.

Емкость C_0 несколько разряжается за время прямого хода. Поскольку при запертом диоде D через конденсаторы C_0 и C протекает один и тот же ток, приращение заряда на них также будет

¹ Конечно, под $I_{к0}$ понимается полный обратный ток коллектора, включая ток термогенерации и ток утечек.

одинаково, а следовательно, напряжений будут обратно пропорциональны емкостям:

$$\Delta U_{C_0} \approx \frac{C}{C_0} U_m. \tag{21-34}$$

Таким образом, в момент $T_{пр}$ компенсирующее напряжение ΔU_a будет иметь значение $KU_m - \Delta U_{C_0}$ вместо KU_m , а это равносильно уменьшению коэффициента передачи повторителя K (см. пунктирные кривые на рис. 21-9, б):

$$K' = K - \frac{\Delta U_{C_0}}{U_m} = K - \frac{C}{C_0}. \tag{21-35}$$

Как видим, емкость C_0 непосредственно влияет на действующий коэффициент передачи K' , а через него на линейность

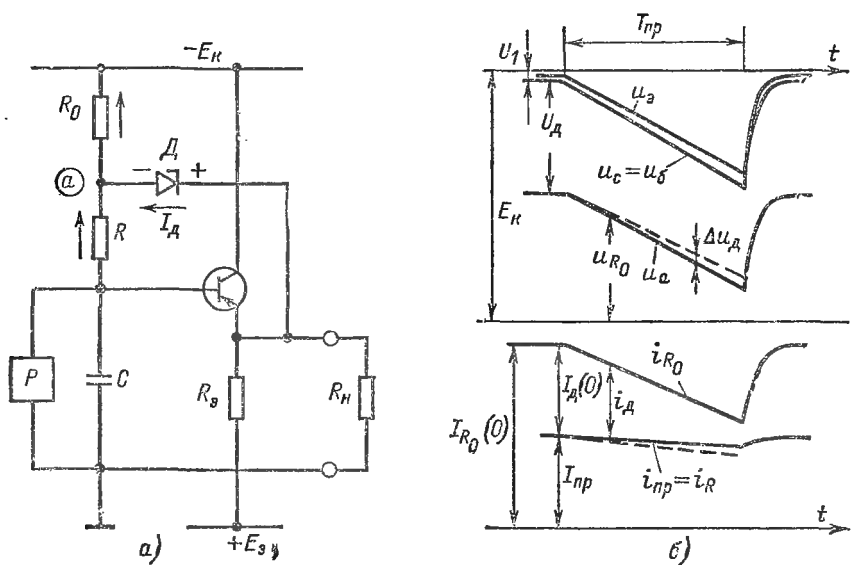


Рис. 21-10. Вариант генератора с повторительной следящей связью.
 а — принципиальная схема; б — временные диаграммы.

пилообразной кривой согласно (21-17). Очевидно, что с точки зрения линейности необходимо достаточно сильно выполнить неравенство $C_0 \gg C$. Однако большое значение C_0 приводит к отрицательным последствиям: затягивается процесс восстановления в конце обратного хода. В самом деле, когда диод D снова откроется, конденсатор C_0 должен будет дозарядиться на величину ΔU_{C_0} через R_d и выходное сопротивление повторителя. Постоянная времени «дозаряджения» пропорциональна емкости C_0 и может значительно превышать постоянную времени CR_k , с которой разряжается рабочий конденсатор. Наиболее радикальным способом уменьшить

время восстановления является замена конденсатора C_0 полупроводниковым стабилитроном (рис. 21-10).

К сожалению, схеме с полупроводниковым стабилитроном тоже свойственны недостатки: во-первых, балластное сопротивление R_0 , обеспечивающее отвод суммарного тока $I_{пр} + I_d$, шунтирует нагрузку по переменному току и, следовательно, уменьшает входное сопротивление повторителя; во-вторых, в начальном состоянии на сопротивлении R_0 падает значительное напряжение и соответственно требуется большее напряжение питания, чем в схеме с конденсатором.

В самом деле, в исходном состоянии имеем:

$$U_a(0) = U_{э2}(0) + U_d \approx U_1 + U_d.$$

Во время прямого хода потенциал $U_{э2}$, а вместе с ним и потенциал U_a возрастают на величину $KU_m \approx U_m$. Следовательно,

$$U_a(T_{пр}) \approx U_1 + U_d + U_m.$$

Разность потенциалов $E_k - U_a$ в интервале $T_{пр}$ уменьшается, а значит, уменьшается ток I_d , поскольку ток $I_{пр}$ остается неизменным. В конце прямого хода ток I_d не должен быть меньше $I_{d, мин}$ — минимального рабочего тока стабилитрона (см. § 3-3). Учитывая это ограничение, можно записать напряжение питания в виде

$$E_k = U_1 + U_d + U_m + (I_{d, мин} + I_{пр}) R_0; \quad (21-36)$$

подставляя E_k в (20-4), получаем коэффициент использования напряжения питания:

$$\xi = \frac{U_m}{U_m + U_1 + U_d + (I_{d, мин} + I_{пр}) R_0}. \quad (21-37)$$

Сумма последних двух слагаемых в знаменателе обычно сравнима с амплитудой U_m или даже превышает ее. Поэтому ГПН со стабилитроном характеризуется небольшим коэффициентом использования (0,5 и менее). Однако этот недостаток искупается малым временем восстановления и отсутствием конденсатора большой емкости.

Если принять $I_{пр} = \text{const}$, то приращения тока в цепи $D-R_0$ будут определяться приращениями потенциала $U_{э2}$. В конце прямого хода, когда $\Delta U_{э2} \approx U_m$, ток стабилитрона уменьшится:

$$\Delta I_d = \frac{U_m}{R_0 + r_d},$$

а напряжение на нем возрастет на величину

$$\Delta U_d = \Delta I_d r_d = \frac{U_m r_d}{R_0 + r_d} \approx \frac{I_{пр} T_{пр}}{C (R_0 / r_d + 1)}, \quad (21-38)$$

где r_d — дифференциальное сопротивление стабилитрона. Соответственно компенсирующее напряжение в конце прямого хода будет $KU_m - \Delta U_d$ вместо KU_m , что равносильно уменьшению коэффициента передачи (см. пунктирные кривые на рис. 21-10, б):

$$K' = K - \frac{\Delta U_d}{U_m} = K - \frac{r_d}{R_0 + r_d}. \quad (21-39)$$

Сравнивая выражения (21-38) и (21-39) с выражениями (21-34) и (21-35), замечаем, что стабилитрон во время прямого хода эквивалентен емкости $C_0 = C(1 + R_0/r_d)$; поэтому должно достаточно сильно выполняться неравенство $R_0 \gg r_d$; такое неравенство одновременно способствует повышению входного сопротивления повторителя (см. выше).

Генераторы с усилительной связью. Типичным примером таких ГПН является интегратор Миллера (рис. 21-11), который относится к генераторам спадающего напряжения с посторонним возбуждением¹.

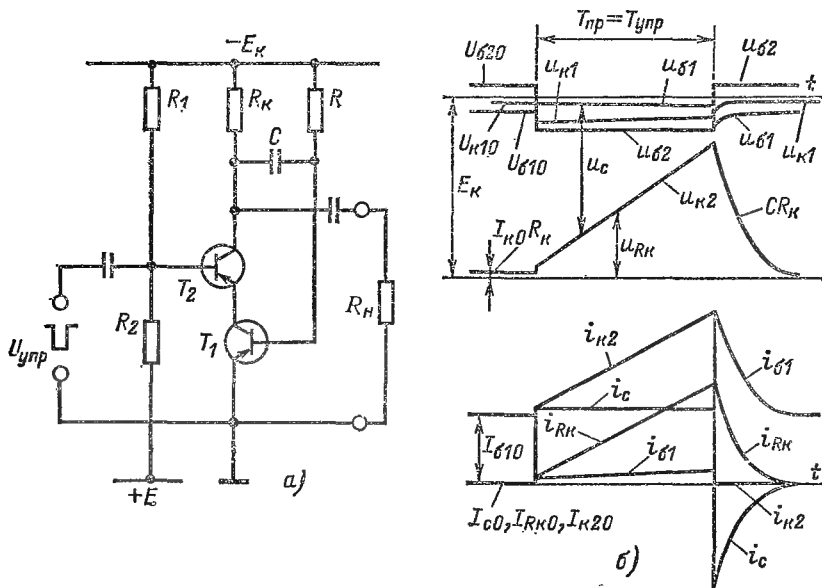


Рис. 21-11. Генератор с усилительной следящей связью.

а — принципиальная схема; б — временные диаграммы.

В начальном состоянии транзистор T_2 заперт положительным смещением на базе. Поэтому ток $I_{э2}$, а значит, и ток $I_{к1}$ практически равны нулю. В то же время в цепи эмиттер — база транзистора T_1 протекает вполне определенный ток

$$I_{б1} = \frac{E_K - U_{б10}}{R} \approx \frac{E_K}{R}. \quad (21-40)$$

В этих условиях T_1 находится в режиме глубокого насыщения; соответственно $U_{к10} = U_{э20} \approx 0$ и, следовательно, запирающее

¹ В ламповой технике интегратор Миллера выполнялся на пентоде, управляемом по третьей сетке. В транзисторном варианте используется последовательное включение двух транзисторов; эта комбинация в какой-то степени аналогична ламповому тетроду.

смещение U_{620} может иметь очень небольшое значение — доли вольта.

Поскольку коллекторный ток отсутствует, напряжение U_{610} можно оценить по «диодной» формуле (2-64б), подставляя $I = I_{61}$:

$$U_{610} \approx \varphi_T \ln \frac{I_{610}}{I_{\text{эо1}}} + I_{610} r_{61}. \quad (21-41)$$

Начальное напряжение на рабочем конденсаторе C в общем виде определяется соотношением

$$U_c(0) = E_k - I_{k02} R_k - U_{610};$$

практически $U_c(0) \approx E_k$.

Отрицательный управляющий импульс $U_{\text{упр}}$ отпирает транзистор T_2 и вызывает появление токов $I_{\text{э2}}$ и I_{k2} . Поскольку входное сопротивление насыщенного транзистора значительно меньше R_k и R_n , ток I_{k2} почти полностью проходит через конденсатор C в базовую цепь транзистора T_1 и а в с т р е ч у начальному току I_{610} . Таким образом, под действием управляющего импульса, во-первых, появляется ток $I_{k1} = I_{\text{э2}}$ и, во-вторых, резко уменьшается ток I_{61} . Оба эти обстоятельства обеспечивают быстрый выход транзистора T_1 из насыщения, так что во время последующего прямого хода он работает в активном режиме. Условие активного режима $U_{k61} < 0$ или $|U_{k\text{э1}}| > U_{\text{э61}}$ легко записать в следующем виде¹:

$$|U_{\text{упр}}| - U_{620} - U_{\text{э62}} > U_{\text{э61}},$$

откуда определяется необходимое значение $U_{\text{упр}}$.

Как показано выше, ток I_{k2} в момент запуска генератора почти полностью идет через конденсатор C и, следовательно, является током прямого хода. Для того чтобы найти этот ток, пренебрежем тепловыми токами и используем соотношения между начальными значениями токов:

$$\begin{aligned} I_{61}(0) &= I_{610} - I_{k2}(0); \\ I_{k2}(0) &= \alpha_2 I_{\text{э2}}(0); \\ I_{\text{э2}}(0) &= I_{k1}(0) = \beta_1 I_{61}(0). \end{aligned}$$

Из этих соотношений нетрудно получить:

$$I_{c \text{ пр}} \approx I_{k2}(0) \approx \frac{\alpha_2 \beta_1}{1 + \alpha_2 \beta_1} I_{610}; \quad (21-42)$$

$$I_{61}(0) = \frac{I_{610}}{1 + \alpha_2 \beta_1}. \quad (21-43)$$

Как видим, $I_{k2}(0) \approx I_{610}$. Ток $I_{\text{э1}}(0)$ мало отличается от $I_{k2}(0)$ и, следовательно, тоже близок к I_{610} . Следовательно, эмиттерный ток, а значит, и напряжение на эмиттерном переходе транзистора T_1 остаются почти неизменными при запуске генератора. В то же время ток базы меняется сильно: $\Delta I_{61} \approx I_{610}$. Значит, на сопротивлении базы получается о м и ч е с к и й с к а ч о к напряжения $\Delta U_{61} \approx I_{610} r_{61}$. Этот скачок обычно составляет десятки доли вольта и передается через конденсатор C на коллектор транзистора T_2 .

¹ Напряжения $U_{\text{э62}}$ и $U_{\text{э61}}$ можно найти либо из входных характеристик, либо из (15-9г). В обоих случаях желательно ориентироваться на максимальные значения этих напряжений при токе $E_k/(R_k \| R_n)$.

Мы подробно остановились на моменте запуска генератора, так как этот этап представляет известные трудности для понимания и не всегда четко описывается в литературе. Последующий этап прямого хода нуждается лишь в небольших пояснениях.

На этапе прямого хода конденсатор разряжается током $I_{cпр}$ и напряжение U_c уменьшается. Приращение ΔU_c можно записать в виде

$$\Delta U_c = -\Delta U_{к2} + \Delta U_{б1},$$

где $\Delta U_{к2} = K_u \Delta U_b$ и $K_u < 0$. Следовательно,

$$\Delta U_{б1} = \frac{\Delta U_c}{1 - K_u}; \quad (21-44a)$$

$$\Delta U_{к2} = \frac{K_u}{1 - K_u} \Delta U_c. \quad (21-44б)$$

Если $|K_u| \gg 1$ (что всегда выполняется), то $\Delta U_{б1} \ll \Delta U_c$ и $\Delta U_{к2} \approx \Delta U_c$, т. е. потенциал $U_{б1}$ почти не меняется, а напряжение $U_{к2}$ «следит» за изменениями напряжения U_c . Таким образом, работа генератора соответствует скелетной схеме на рис. 21-8, б, если положить $I_{пр} = I_{б10}$; $E = E_k$; $U_1 = U_c(0) \approx E_k$ и $k = |K_u|/(1 + |K_u|)$.

Коэффициент нелинейности рассчитывается по формулам (21-17), а поправочный коэффициент ϑ — по формуле (21-24в). Поскольку $1 - k \approx 1/|K_u|$, получаем:

$$\varepsilon \approx \frac{U_m}{|K_u| E_k}, \quad (21-45)$$

где коэффициент усиления K_u можно найти из (7-21), полагая $R_n = \infty$. Коэффициент K_u возрастает с увеличением R_k и обычно составляет 50—100. При этом коэффициент нелинейности $\varepsilon = 0,5 \div 1\%$.

Большие значения R_k , выгодные с точки зрения коэффициента усиления, оказываются невыгодными с точки зрения длительности обратного хода. Действительно, резистор R_k является пассивным зарядным элементом: через него конденсатор C заряжается до напряжения $U_c(0)$ по окончании сигнала. Заряд происходит с постоянной времени, близкой к CR_k .

Минимальное значение выходного потенциала (U_2) определяется напряжением $U_{упр}$, так как потенциал $U_{к2}$ уменьшается (по модулю) лишь до тех пор, пока не станет равным потенциалу $U_{б2} = U_{упр} - U_{б20}$. После этого наступает насыщение транзистора T_2 и прямой ход заканчивается. Следовательно (см. рис. 21-1, г):

$$U_{2\min} = U_{упр} - U_{б20};$$

$$U_1 \approx E_k,$$

откуда максимальная амплитуда выходного напряжения

$$U_{m\max} \approx E_k - U_{упр} - U_{б20}. \quad (21-46)$$

Фантастроны. В основе этого типа ГПН лежит тот же интегратор Миллера. Особенности фантастронов заключаются в следующем. Во-первых, длительность прямого хода не зависит от длительности управляющего сигнала, а определяется внутренними процессами в схеме (при этом управляющий сигнал представляет собой короткий пусковой импульс). Во-вторых, амплитуда пилообразного напряжения, определяющая длительность прямого хода, легко регулируется изменением напряжения смещения, что важно для ряда применений [162]. Первая особенность заставляет отнести

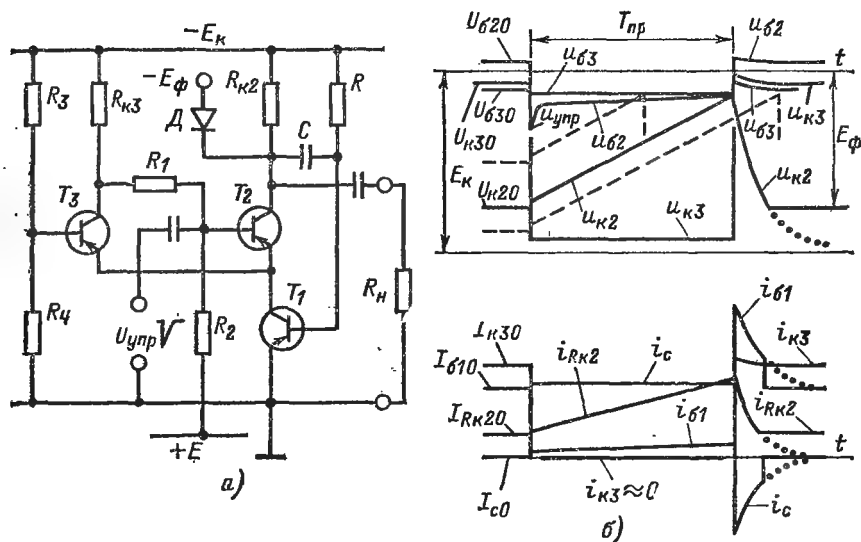


Рис. 21-12. Фантастрон с усилительной следящей связью.

а — принципиальная схема; б — временные диаграммы.

фантастроны к ГПН с самовозбуждением¹, а тот факт, что в их основе лежит интегратор Миллера, говорит о генерации с п а д а ю щ е г о напряжения.

Для того чтобы конец прямого хода определялся в н у т р е н н и м и процессами, в схеме должно быть пороговое устройство. В ламповой технике для этого использовался *транзитронный эффект* (перераспределение токов анода и экранирующей сетки) в пентоде [164, 168]. В транзисторной технике перераспределение токов приходится обеспечивать с помощью дополнительного транзистора. Типичная схема фантастрона показана на рис. 21-12.

Функции транзисторов T_1 и T_2 те же, что и в схеме на рис. 21-11, а транзисторы T_2 и T_3 образуют своеобразный триггер с эмиттерной

¹ Помимо ждущего режима фантастроны могут работать и в режиме автоколебаний [168, 169].

связью, в котором роль резистора R_3 играет транзистор T_1 . Во время прямого хода транзистор T_3 и диод D заперты, так что формирование пилообразного напряжения происходит в тех же условиях, что и в ГПН на рис. 21-11. Единственное, но существенное отличие состоит в том, что потенциал U_{62} в фантастроне определяется не управляющим импульсом (который имеет очень малую длительность и заканчивается в самом начале прямого хода), а делителем $R_{к3}$, R_1 , R_2 , питаемым от напряжений E_k и E . Поэтому заключительная стадия прямого хода протекает следующим образом.

Когда потенциал $U_{к2}$, уменьшаясь, делается равным потенциалу U_{62} , транзистор T_2 попадает в режим насыщения, т. е. «стягивается в точку», и нагрузкой каскада T_1 становится суммарное сопротивление $R_2 \parallel (R_{к3} + R_1) \parallel R_{к2} \parallel R_n$ вместо прежнего $R_{к2} \parallel R_n$. При этом несколько уменьшается коэффициент усиления K_u и соответственно увеличивается коэффициент нелинейности (21-45). Однако качественное изменение процесса (окончание прямого хода) наступает лишь тогда, когда потенциал U_{63} , уменьшаясь вместе с потенциалом транзистора T_2 , делается равным потенциалу U_{63} (последний задается делителем R_3 , R_4). В этот момент открывается транзистор T_3 и нарастающий ток $I_{к3}$, ответвляясь через R_1 , вызывает регенерацию в системе T_2 — T_3 . Результатом регенерации является запираание транзистора T_2 , сопровождающееся сильным возрастанием тока I_{61} , как и в схеме на рис. 21-11. Большой ток I_{61} приводит к насыщению транзисторов T_1 и T_3 . Режим насыщения сохраняется и тогда, когда ток I_{61} падает до начального значения $I_{610} \approx E_k/R$, поскольку и это значение намного больше, чем ток до регенерации. Насыщение транзистора T_2 почти совпадает с отпиранием T_3 , поэтому некоторые авторы [169] считают насыщение транзистора T_2 причиной регенерации, что, конечно, неверно.

После запираания транзистора T_2 наступает обратный ход: конденсатор C заряжается через резистор $R_{к2}$ до тех пор, пока потенциал $U_{к2}$ не достигнет значения E_ϕ . В этот момент отпирается диод D и фиксирует потенциал $U_{к2}$ на уровне $E_\phi < E_k$. Таким образом, как видно из рис. 21-12, б, пилообразное напряжение меняется в пределах от $U_1 \approx E_\phi$ [точнее, от $E_\phi - \Delta U_{к2}$, где начальный скачок $\Delta U_{к2}$ рассчитывается по формуле (21-44)] до $U_2 \approx U_{63} + U_{632}(T_{пр})$, где U_{63} — потенциал, задаваемый делителем R_3 , R_4 во время прямого хода (когда T_3 заперт). *Меняя потенциал E_ϕ , можно плавно менять амплитуду ($U_m \approx E_\phi - U_{63}$), а значит, и длительность прямого хода ($T_{пр} = U_m / U'_c$), обеспечивая при этом высокую линейность зависимости $T_{пр}(E_\phi)$. Фиксация коллекторного потенциала, как и в других аналогичных случаях, приводит к сокращению времени восстановления, в данном случае — времени обратного хода.*

Из приведенного описания следует, что при расчете фантастрона нужно обеспечить в исходном состоянии насыщение транзисторов T_3 , T_1 и запираание транзистора T_2 , а в рабочем режиме — насыщение транзистора T_2 при отпирании

T_3 (когда $U_{эб3} = 0$). Для этого должны выполняться условия

$$\beta_3 I_{б30} > I_{к30}; \quad (21-47a)$$

$$\beta_1 I_{б10} > I_{к10} = I_{к30} + I_{б30}; \quad (21-47б)$$

$$U_{б20} > U_{к10} + n\varphi_T \quad (21-47в)$$

($n = 3 \div 4$ при глубокой отсечке);

$$U_{к2}(T_{пр}) = U_{б2}(T_{пр}) \approx U_{б3} + U_{эб2}(T_{пр}). \quad (21-47г)$$

Потенциал $U_{б3}$ и близкий к нему потенциал $U_{к2}(T_{пр})$ выбираются достаточно малыми, так как увеличение их приводит к уменьшению выходной амплитуды. Однако значение $U_{б3}$ должно превышать напряжение $U_{эб1}(T_{пр})$, с тем чтобы в момент отпираания транзистора T_3 транзистор T_1 еще находился в активном режиме. В противном случае прямой ход закончится насыщением транзистора T_1 , а транзистор T_3 не откроется; соответственно не запрется T_2 и ГПН не вернется в исходное состояние.

Помимо условий (21-47) важно еще обеспечить достаточную **с т а б и л ь н о с т ь** потенциала $U_{б3}$, например, по отношению к изменениям тока $I_{к0}$ (или тока утечки), протекающего в базе запертого транзистора T_3 . Нестабильность величины $U_{б3}$ означает нестабильность амплитуды, а вместе с нею и длительности прямого хода.

Глава двадцать вторая

ПРЕОБРАЗОВАТЕЛИ ПОСТОЯННОГО НАПРЯЖЕНИЯ

22-1. ПРИНЦИП ДЕЙСТВИЯ И КЛАССИФИКАЦИЯ

При разработке электронной аппаратуры, особенно переносной, часто требуется получить достаточно высокое напряжение постоянного тока (от десятков до тысяч вольт) с помощью низковольтных источников питания, например аккумуляторных или гальванических батарей (с напряжением от 1—2 до 25—50 В). Для этого используются так называемые преобразователи постоянного напряжения (ППН).

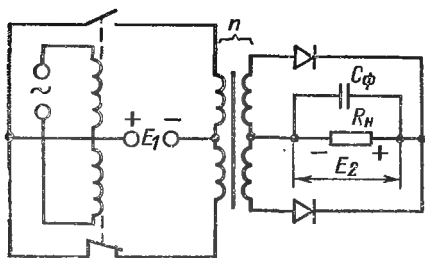


Рис. 22-1. Принципиальная схема вибропреобразователя.

До появления транзисторов основным типом ППН был вибропреобразователь, схема которого показана на рис. 22-1 в двухтактном варианте. Здесь половинки первичной обмотки трансформатора с помощью механических контактов поочередно подключаются к источнику питания E_1 . В результате на вторичной (повышающей) обмотке

получается симметричное прямоугольное напряжение с амплитудой nE_1 , которое затем выпрямляется и фильтруется. Таким образом, принцип действия ППН заключается в преобразовании низкого постоянного напряжения в переменное, повышении последнего с помощью трансформатора и последующем выпрямлении высокого напряжения. Разумеется, любой ППН может быть выполнен и по одноконтурной схеме. Однако последняя обладает по сравнению с двухтактной теми же недостатками, что и однополупериодный выпрямитель по сравнению с двухполупериодным. Поэтому двухтактные схемы предпочтительнее.

Использование механических контактов на первом этапе преобразования приводит к плохой надежности, малому сроку службы, а также к ограничению рабочей частоты, которую желательно иметь

большой для облегчения фильтра¹. Эти трудности в значительной мере устраняются, если вместо механических контактов использовать транзисторные ключи. При этом, вообще говоря, отпадает необходимость в специальном генераторе для управления ключами, так как можно осуществить положительную обратную связь между ключевыми транзисторами, которая обеспечивает их поочередное отпирание и запираение.

Скелетная схема ППН показана на рис. 22-2. Как видим, ППН состоит из релаксационного генератора $РГ$, ключа K , трансформатора $Тр$ и нагруженного выпрямителя B^* . Такая общая схема может быть реализована в виде трех основных вариантов:

1) ППН с автономным генератором; 2) ППН с бестрансформаторной обратной связью; 3) ППН с трансформаторной обратной связью.

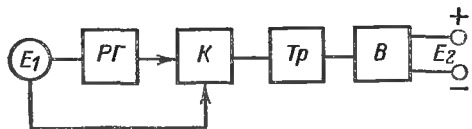


Рис. 22-2. Скелетная схема преобразователей постоянного напряжения.

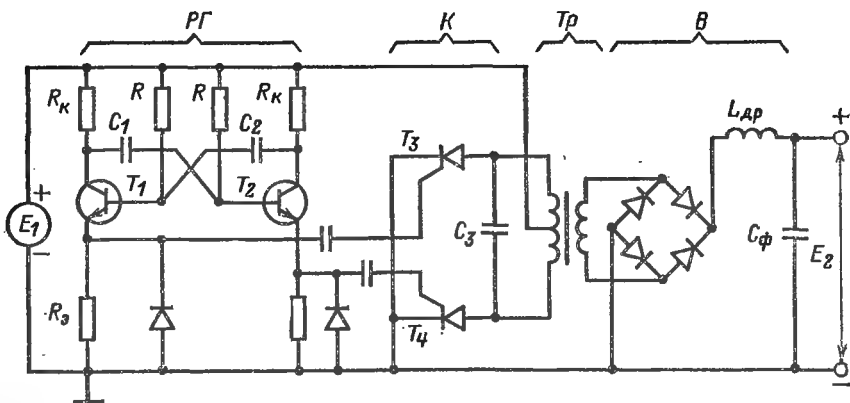


Рис. 22-3. Преобразователь постоянного напряжения с автономным генератором и управляемыми ключами.

В первой группе ППН ключами управляет специальный генератор, который питается от источника преобразуемого напряжения E_1 . Пример такого ППН показан на рис. 22-3. Здесь транзисторы T_1 и T_2 включены по схеме мультивибратора, а управляемыми ключами являются тиристоры T_3 и T_4 . Выходные импульсы мультивибратора

¹ Емкости конденсаторов и индуктивности дросселей фильтра обратно пропорциональны частоте.

* Вообще говоря, вместо релаксационного генератора можно использовать генератор синусоидальных колебаний; однако при этом ухудшаются основные показатели ППН: уменьшается к. п. д., падает выходная мощность, растет коэффициент пульсаций на выходе фильтра.

снимаются не с коллекторов, а с эмиттеров, чтобы изолировать времязадающие цепи от нагрузки — базовых цепей тиристоров. Гашение тиристоров обеспечивается конденсатором C_3 , передающим отрицательный перепад напряжения с коллектора отпирающего ключа на коллектор открытого, который нужно запереть. Следовательно, гашение не входит в функцию мультивибратора. Это позволяет строить последний на сравнительно маломощных транзисторах.

Во второй группе ППН ключевые транзисторы входят в схему релаксационного генератора, но обратные связи осуществляются

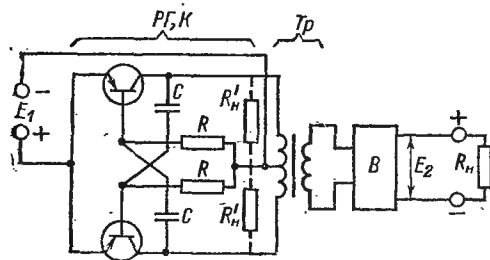


Рис. 22-4. Преобразователь постоянного напряжения с совмещенными генератором и ключевой схемой.

без дополнительных обмоток трансформатора. Пример такого ППН показан на рис. 22-4, где транзисторы включены по схеме мультивибратора [172].

Специфика этого мультивибратора по сравнению со схемой на рис. 18-1 заключается в том, что роль коллекторных резисторов R_k выполняет пересчитанное сопротивление нагрузки R'_n . Поэтому в случае

переменной нагрузки при выборе времязадающего сопротивления R нужно подставлять в неравенства (18-1) минимальное значение R'_n . В остальном мультивибратор является независимым узлом и расчет схемы в целом не представляет затруднений.

В третьей группе ППН трансформатор имеет специальные обмотки обратной связи и поэтому является органической частью релаксационной схемы. Такие ППН, имеющие наибольшее распространение на практике [173, 174], подробнее рассмотрены ниже.

22-2. АНАЛИЗ ПРЕОБРАЗОВАТЕЛЯ С ТРАНСФОРМАТОРНОЙ ОБРАТНОЙ СВЯЗЬЮ

Схема простейшего ППН такого типа показана на рис. 22-5; нагрузкой для простоты считается линейное сопротивление R_n . Такая схема является по существу двухтактным блокинг-генератором без времязадающей цепи RC (ср. с рис. 20-1, а).

Как следует из выражения (20-23), отсутствие конденсатора C (т. е. условие $C = \infty$ или $\tau_c = \infty$) не является препятствием для формирования импульсов: длительность их в этом случае будет определяться постоянной времени τ_L . Что касается паузы, то она обеспечивается соответствующей полярностью напряжений на обмотках обратной связи. В самом деле, когда насыщен транзистор T_1 , э. д. с. E_1 оказывается приложенной к обмотке w_{k1} и полярности напряжений на обмотках будут такими, как показано на рис. 22-5 без скобок. При этом транзистор T_2 заперт положительным смеще-

нием на базе. Вследующем полупериоде, когда насыщен транзистор T_2 , э. д. с. E_1 оказывается приложенной к обмотке $w_{к2}$ и полярности напряжений на обмотках будут обратными (на рис. 22-5 показаны в скобках). Соответственно запертым окажется транзистор T_1 . Таким образом, пауза в одной половине схемы соответствует импульсу в другой половине.

Для анализа ППН в принципе можно использовать теорию формирования вершины импульса в блокинг-генераторе (§ 20-4). В частности, эквивалентная схема ППН (рис. 22-6) получается из эквивалентной схемы, показанной на рис. 20-6, путем исключения конденсатора C' . Однако в теории блокинг-генератора импульсы считаются одиночными, т. е. подразумевается, что к началу очередного импульса выброс от предыдущего заканчивается, тогда как в ППН импульсы следуют непосредственно друг за другом и выброс от каждого из них (обусловленный остаточным током намагничивания) влияет на длительность следующего импульса. Кроме того, в ППН в отличие от блокинг-генераторов обычно стараются выполнить условие $R'_H < r'_6$ с тем, чтобы мощность в основном рассеивалась в нагрузке и тем самым к. п. д. был достаточно высок. Конечно, указанные особенности можно было бы учесть в рамках теории, изложенной в § 20-4, но проще проанализировать ППН отдельно.

Начнем анализ с момента включения э. д. с. E_1 . В результате флуктуаций, а также неизбежной асимметрии схемы ток в одной из ее половин будет нарастать несколько быстрее, чем в другой. Если выполняется условие регенерации (20-10), процесс получается лавинообразным и один из транзисторов окажется насыщенным, а другой — запертым. Пусть насыщен транзистор T_1 . Тогда, пренебрегая сопротивлением участка коллектор — эмиттер, получаем для коллекторной цепи эквивалентную схему, показанную на рис. 22-6, где r'_6 и R'_H — приведенные сопротивления базы и нагрузки¹.

Из схемы видно, что ток базы не меняется и равен:

$$I_6 = \frac{1}{n_6} \frac{E_1}{r'_6}, \quad (22-1)$$

где $n_6 = w_6/w_k$ — коэффициент трансформации. Как известно (см. § 15-5), заряд в базе при заданном токе I_6 нарастает с постоянной времени τ , стремясь к установившемуся значению $I_6\tau$. На прак-

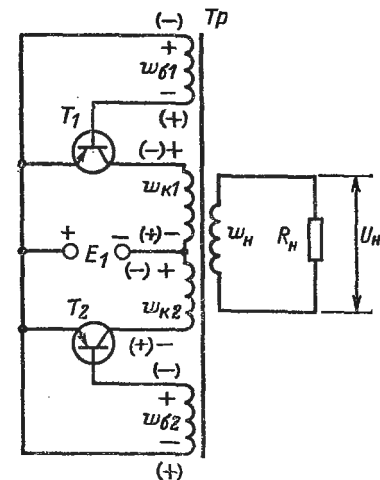


Рис. 22-5. Простейший преобразователь постоянного напряжения с трансформаторной обратной связью.

¹ Генератор тока $I_{\mu m}$ пока что учитывать не следует; его происхождение будет пояснено ниже.

тике в большинстве ППН выполняется неравенство $T_n \gg \tau$, и, следовательно, установившееся значение заряда будет достигнуто задолго до окончания импульса:

$$Q(T_n) = I_6(\tau). \quad (22-2)$$

Граничный заряд $Q_{гр}$ согласно (15-27) пропорционален току I_k и в момент T_n равен:

$$Q_{гр}(T_n) = \frac{\tau}{\beta} I_k(T_n). \quad (22-3)$$

Приравняв (22-2) и (22-3), получим уравнение для длительности импульса, которое по форме совпадает с критерием насыщения (15-7):

$$I_k(T_n) = \beta I_6. \quad (22-4)$$

Из рис. 22-6 следует, что коллекторный ток состоит из постоянной составляющей $I'_6 + I'_n$ и переменной составляющей I_{μ} — тока намагничивания сердечника.

Поскольку $I_{\mu}(s) = E_1/sL$, получаем:

$$i_{\mu}(t) = \frac{E_1}{L} t, \quad (22-5)$$

т. е. ток намагничивания, а значит, и коллекторный ток меняются линейно со скоростью E_1/L (рис. 22-7).

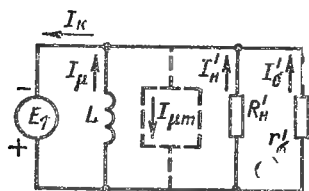


Рис. 22-6. Эквивалентная схема преобразователя постоянного напряжения во время формирования импульсов (генератор тока I_{μ} действует, начиная со второго импульса).

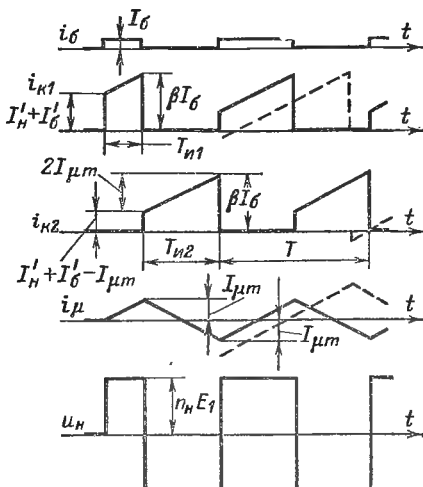


Рис. 22-7. Временные диаграммы преобразователя постоянного напряжения для первых трех импульсов (пунктиром показан случай $I_{\mu m} > I'_6 + I'_n$).

За время первого импульса коллекторный ток должен возрасти от начального значения $I'_6 + I'_n$ до значения βI_6 , определяемого выражением (22-4). На такую же величину возрастет и ток намагничивания I_{μ} . Обозначим это приращение через $I_{\mu m}$ и запишем его в следующем виде:

$$I_{\mu m} = \beta I_6 - (I'_6 + I'_n) = E_1 \left[(\beta' - 1) \frac{1}{r'_6} - \frac{1}{R'_H} \right], \quad (22-6)$$

где $\beta' = \beta/n_6$ — приведенный коэффициент усиления [см. (20-23)].

В момент $T_{н1}$ транзистор T_1 выходит из насыщения и наступает регенеративный процесс, в результате которого транзистор T_1 запирается, а транзистор T_2 попадает в режим насыщения. В течение короткого регенеративного процесса ток намагничивания не успевает существенно измениться, сохраняя значение $I_{\mu m}$ и прежнее направление. Поэтому в начале второго импульса через обмотку $\omega_{к2}$ помимо тока $I'_6 + I'_H$, направленного из коллектора, протекает ток $I_{\mu m}$, направленный в обратную сторону — в сторону коллектора. Этот ток, характеризующий остаточную намагниченность сердечника, отражен на рис. 22-6 генератором тока $I_{\mu m}$, показанным пунктиром. Нетрудно догадаться, что длительность второго импульса будет вдвое больше, чем первого, так как коллекторный ток (при той же скорости нарастания) должен теперь измениться на $2I_{\mu m}$, чтобы достигнуть уровня $\beta I'_6$ (см. временные диаграммы на рис. 22-7).

В конце второго импульса ток намагничивания будет иметь такую же величину $I_{\mu m}$, как и в конце первого. Поэтому, начиная со второго импульса, процессы в ППН устанавливаются и длительность импульсов T_n определяется удвоенной амплитудой тока намагничивания ($2I_{\mu m}$) и скоростью его изменения (E_1/L). Считая, что в симметричной схеме полный период $T = 2T_n$, получаем:

$$T = \frac{4I_{\mu m}L}{E_1}. \quad (22-7)$$

Подставив сюда выражение (22-6) для тока $I_{\mu m}$, представим период и рабочую частоту F в следующем виде:

$$T = \frac{1}{F} = 4L \left[(\beta' - 1) \frac{1}{r'_6} - \frac{1}{R'_H} \right]. \quad (22-8)$$

Из формулы (22-8) следует, что рабочая частота обратно пропорциональна индуктивности коллекторной обмотки и не зависит от преобразуемой э. д. с. E_1 . Частота увеличивается с увеличением тока нагрузки, т. е. с уменьшением сопротивления R'_H ; эта зависимость существенна тогда, когда значение R'_H сравнимо со значением r'_6/β' , т. е. когда $R'_H \ll r'_6$.

Очевидно, что правая часть выражения (22-8) должна быть положительной, т. е. должно выполняться условие

$$\beta' > 1 + \frac{r'_6}{R'_H}. \quad (22-9)$$

Это условие, являющееся частным случаем неравенства (20-23) при $\tau_c = \infty$ и $\tau = 0$, накладывает определенные ограничения на параметры схемы. А именно, учитывая, что $r'_6 = r_6/n_6^2$, получаем:

$$n_6^2 - \beta n_6 + \frac{r_6}{R'_H} < 0. \quad (22-10)$$

Такое неравенство выполняется только при достаточно большом коэффициенте передачи β :

$$\beta > 2\sqrt{r_6/R'_H}. \quad (22-11)$$

Соотношения (22-11) должно соблюдаться при минимальном значении R'_H и максимальном значении r'_6 и является одним из критериев при выборе транзисторов. Если коэффициент β превышает правую часть (22-11) более чем в 2 раза¹, то корни левой части (22-10) имеют приближенные значения β и $r'_6/\beta R'_H$. Коэффициент n_6 должен лежать в интервале между этими корнями. Обычно значения $n_6 > 1$ неприемлемы с точки зрения к. п. д. схемы (поскольку при $n_6 > 1$ имеем $r'_6 < r_6$ и может оказаться $r'_6 \approx R'_H$, т. е. «паразитная» мощность P_{r_6} будет сравнима с полезной мощностью P_{R_H}). Поэтому запишем только то ограничение, которое накладывается на величину n_6 с н и з у:

$$n_6 > r_6/\beta R'_H. \quad (22-12)$$

Неравенство (22-12) лежит в основе выбора коэффициента трансформации при расчете ППН.

Весь проведенный анализ был основан на постоянстве индуктивности. Между тем с ростом тока намагничения индукция в сердечнике трансформатора может достигнуть таких значений, при которых магнитная проницаемость и соответственно индуктивность L заметно уменьшаются (насыщение сердечника). В этом случае, как видно из выражения (22-5), скорость нарастания коллекторного тока увеличится и равенство (22-4) будет достигнуто раньше, чем при постоянном значении L (рис. 22-8). Иначе говоря, насыщение сердечника приводит к уменьшению периода и увеличению рабочей частоты по сравнению с расчетными значениями. Кроме того, поскольку ток намагничения согласно (22-6) пропорционален э. д. с. E_1 , изменения последней будут влиять на степень насыщения сердечника и тем самым на рабочую частоту. Строгий анализ этих явлений, разумеется, сложен. Его можно упростить, аппроксимируя кривую намагничения ломаной линией [175].

Рис. 22-8. Случай насыщения сердечника вблизи максимальных значений тока намагничения (пунктирная кривая соответствует отсутствию насыщения).

На практике стараются избегать насыщения, выбирая достаточно большое сечение сердечника. Для этого нужно воспользоваться известными выражениями

$$L = (\mu_0 \mu) \frac{\omega_k^2 S}{l_{cp}}; \quad (22-13a)$$

$$\Psi_m = LI_{\mu m} = \omega_k B_m S \cdot 10^{-8}, \quad (22-136)$$

где μ_0 — магнитная постоянная (см. табл. 1-2); μ — относительная магнитная проницаемость сердечника; l_{cp} — средняя длина магнит-

¹ Такое превышение следует всегда обеспечивать на практике, учитывая, что исходное неравенство (22-9) получается в результате пренебрежения величиной τ/τ_L в более точном неравенстве (20-23).

ной линии в сердечнике; S — площадь сечения сердечника; Ψ_m — максимальное потокоцепление; B_m — максимальная индукция, Гс. Подставляя значение L из (22-13а) в (22-13б), получаем:

$$\frac{w_k}{I_{cp}} = \frac{B_m \cdot 10^{-8}}{(\mu_0 \mu) I_{\mu m}}. \quad (22-14)$$

Подставляя произведение $LI_{\mu m}$ из (22-7) в (22-13б), получаем:

$$S = \frac{E_1 T \cdot 10^8}{4 w_k B_m}. \quad (22-15)$$

Выражения (22-14) и (22-15) используются при расчете следующим образом. Задаваясь размером сердечника, находим тем самым значение I_{cp} . Потом, задаваясь достаточно малым значением B_m , из (22-14) определяем w_k , а из (22-15) — площадь S .

Регенеративные процессы в ППН практически не отличаются от аналогичных процессов в блокинг-генераторе. Поэтому длительность фронтов, если она представляет интерес, можно оценивать по формуле (20-12), а оптимальный коэффициент трансформации — по формуле (20-13).

22-3. ОСОБЕННОСТИ ПРАКТИЧЕСКОЙ СХЕМЫ

Простейший ППН (рис. 22-5) является вполне работоспособным устройством. Однако в ряде случаев он обладает такими недостатками, для устранения которых в схему приходится вводить дополнительные элементы. Рассмотрим причины этих осложнений.

Поскольку для рассматриваемого ППН характерно условие $T/2 \gg \tau$, нельзя не считаться с модуляцией сопротивления базы во время импульса (см. конец § 20-4). Из формулы (22-8) ясно, что уменьшение сопротивления r_6 (при высоком уровне инжекции) приведет к уменьшению рабочей частоты по сравнению с ее расчетным значением, полученным исходя из немодулированной величины r_6^* . Учет модуляции, как и в случае блокинг-генератора, неоправданно сложен. Поэтому, если желательно обеспечить достаточно строгий расчет схемы, следует включить последовательно с базами транзисторов резистор R_6 с сопротивлением $R_6 > r_{60}$ (рис. 22-9), где r_{60} — немодулированное сопротивление базы. При этом весь проведенный анализ остается в силе, но под величиной r_6 нужно понимать сумму $r_{60} + R_6$. Наличие резистора R_6 уменьшает влияние разброса параметров транзисторов, а также расход мощности в базовых цепях.

Поскольку длительность фронтов согласно (20-12) увеличивается при наличии внешнего резистора R_6 , последний обычно шун-

* Этот вывод качественно правилен. Однако следует иметь в виду, что формула (22-8) получена из условия неизменного тока базы. Модуляция сопротивления r_6 приводит к увеличению тока и формула (22-8) оказывается неточной.

тируют конденсатором C_6 , который «закорачивает» резистор на время переходного процесса. Величина емкости C_6 должна удовлетворять условиям

$$t_{\phi} < C_6 r_{6. \min} \ll \frac{T}{2}, \quad (22-16)$$

в которых под $r_{6. \min}$ следует понимать модулированное сопротивление базы.

При подключении ППН к источнику E_1 схема часто не самовозбуждается, так как начальные токи транзисторов малы и величина β недостаточна для выполнения условий регенерации. Чтобы обеспечить самовозбуждение, базу одного из транзисторов присоединяют к источнику E_1 через резистор (R_1 на рис. 22-9). Этот резистор задает такой базовый ток, при котором коэффициент β удовлетворяет неравенствам (20-10).

Нагрузкой преобразователя мы считали выше активное линейное сопротивление R_n . В случае д в у х т а к т н о г о ППН с двух-

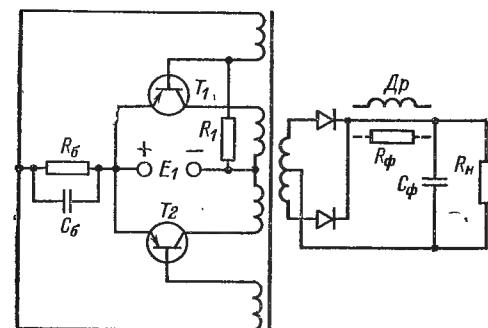


Рис. 22-9. Практическая схема преобразователя постоянного напряжения вместе с выпрямителем (дроссель Dp в двухтактной схеме необязателен).

полупериодным выпрямителем это вполне оправдано, так как напряжение на выходе выпрямителя в течение обоих полупериодов равно $n_n E_1$ и претерпевает небольшие изменения только во время коротких фронтов. Эти изменения сглаживаются емкостью фильтра, которая, учитывая малую длительность фронтов, может иметь сравнительно небольшое значение. Под сопротивлением R_n , строго говоря, следует понимать

сумму $R_n + R_d$, где R_d — сопротивление выпрямительного диода или (в случае мостовой схемы) двух выпрямительных диодов.

В од н о т а к т н о м ППН разряд емкости фильтра C_{ϕ} за время паузы может быть существенным и тогда на эквивалентной схеме (см. рис. 22-6) следует резистор R'_n зашунтировать конденсатором C_{ϕ} . Начальное напряжение на нем (которое можно отразить с помощью э. д. с., включенной последовательно с C_{ϕ}) меньше, чем E_1 , поэтому в начальный момент получается большой скачок тока I_n , ограниченный только сопротивлениями диода, обмоток трансформатора и участка коллектор — эмиттер транзистора. Для предотвращения таких скачков тока, которые могут вывести из строя транзистор, целесообразно использовать фильтр, у которого дроссель пред ш е с т в у е т конденсатору C_{ϕ} .

СТАБИЛИЗАТОРЫ НАПРЯЖЕНИЯ

23-1. ВВЕДЕНИЕ

Полупроводниковые стабилизаторы напряжения используются в основном для питания низковольтной аппаратуры (в первую очередь транзисторной). Как правило, выходные напряжения полупроводниковых стабилизаторов не превышают 50 В.

При разработке стабилизаторов нужно обеспечить две группы показателей: 1) максимальное (оно же номинальное) выходное напряжение $U_{2\text{макс}}$, диапазон его регулирования и его допустимую относительную нестабильность $\delta_2 = \Delta U_{2\text{доп}}/U_2$; 2) максимальный (он же номинальный) ток нагрузки $I_{н.\text{макс}}$ и диапазон его изменений ΔI_n . Как правило, принимают $\Delta I_n = I_{н.\text{макс}}$, т. е. считают $I_{н.\text{мин}} = 0$; стабилизатор, рассчитанный на $I_{н.\text{мин}} \neq 0$, может выйти из строя при случайном сбросе нагрузки или в момент включения (если нагрузка имеет индуктивную составляющую).

Для того чтобы нестабильность выходного напряжения укладывалась в заданные пределы как при изменениях нагрузочного тока, так и при изменениях питающего напряжения, стабилизатор должен иметь соответствующие значения дифференциальных параметров — выходного сопротивления и коэффициента стабилизации. К этим необходимым параметрам обычно добавляют желательный к. п. д. (особенно в случае мощных стабилизаторов) и допустимый дрейф выходного напряжения.

Приведем определения перечисленных параметров, воспользовавшись обозначениями на рис. 23-1.

Выходное сопротивление характеризует изменение выходного напряжения при изменении тока нагрузки и постоянном входном напряжении ¹:

$$R_{\text{вых}} = - \left. \frac{dU_2}{dI_n} \right|_{U_1 = \text{const}} \quad (23-1a)$$

Коэффициент стабилизации характеризует «процентное» изменение входного (питающего) напряжения по отношению к «процентному» изменению выходного напряжения при постоянной нагрузке:

$$K = \frac{U_2}{U_1} \left. \frac{dU_1}{dU_2} \right|_{R_n = \text{const}} \quad (23-16)$$

¹ Практически при расчете и измерении сопротивления $R_{\text{вых}}$ задают на выходе приращение dU_2 и делят его на соответствующее приращение $dI_n = -dI_n$.

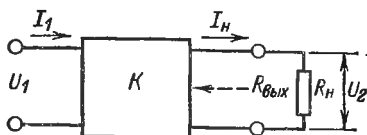


Рис. 23-1. Обозначение электрических величин на схемах стабилизаторов.

Коэффициент полезного действия стабилизатора есть отношение номинальной мощности в нагрузке к номинальной входной мощности:

$$\eta = \frac{U_{2\text{ макс}} I_{\text{н. макс}}}{U_{1\text{ ном}} I_{1\text{ ном}}} \quad (23-2)$$

Временной и температурный дрейф характеризуют абсолютными или относительными изменениями выходного напряжения за определенный промежуток времени или в определенном интервале температур.

При расчете стабилизаторов номинальное входное напряжение непосредственно не задается, но задаются его относительные допуски — верхний и нижний ($+\delta_{в1}$ и $-\delta_{н1}$). Тогда величины

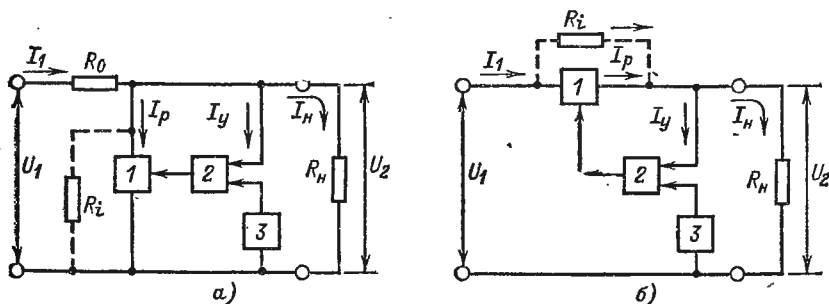


Рис. 23-2. Скелетные схемы стабилизаторов. 1

а — параллельного типа; б — последовательного типа; 1 — регулирующий элемент; 2 — сравнивающий и усилительный элемент; 3 — опорный элемент.

$U_{1\text{ мин}}$, $U_{1\text{ ном}}$, $U_{1\text{ макс}}$ и $(\Delta U_1)_{\text{ макс}}$ будут связаны следующими соотношениями:

$$U_{1\text{ мин}} = U_{1\text{ ном}} (1 - \delta_{н1}); \quad (23-3a)$$

$$U_{1\text{ макс}} = U_{1\text{ ном}} (1 + \delta_{в1}); \quad (23-3б)$$

$$(\Delta U_1)_{\text{ макс}} = U_{1\text{ ном}} (\delta_{в1} + \delta_{н1}) = U_{1\text{ ном}} \delta_1. \quad (23-3в)$$

Определения (23-1) позволяют ввести понятия частных нестабильностей выходного напряжения по току и напряжению. Для этого общую относительную нестабильность δ_2 запишем как сумму

$$\delta_1 = \delta_2^I + \delta_2^U. \quad (23-4a)$$

Согласно (22-1a)

$$\delta_2^I = \frac{|\Delta U_2^I|}{U_2} = \frac{\Delta I_{\text{н}} R_{\text{вых}}}{U_2} = \frac{I_{\text{н. макс}} R_{\text{вых}}}{U_2} = \frac{R_{\text{вых}}}{R_{\text{н. мин}}}. \quad (23-4б)$$

Согласно (23-1б)

$$\delta_2^U = \frac{\Delta U_2^U}{U_2} = \frac{\Delta U_1}{K U_1} = \frac{\delta_1}{K}. \quad (23-4в)$$

Выражение (23-4в) позволяет оценить необходимое значение коэффициента стабилизации, а выражение (23-4б) — необходимое значение выходного сопротивления. Если выходное сопротивление нельзя сделать достаточно малым, то из (23-4б) легко найти максимально допустимый ток нагрузки. Очевидно, что допустимый ток нагрузки и желательное выходное сопротивление находятся в обратной зависимости друг от друга.

Стабилизаторы напряжения делятся на два основных класса: параллельного типа (параллельные) и последовательного типа (последовательные). Скелетные схемы того и другого типов показаны на рис. 23-2.

23-2. СТАБИЛИЗАТОРЫ ПАРАЛЛЕЛЬНОГО ТИПА

Общие свойства. Если пренебречь током через R_i (внутреннее сопротивление регулирующего элемента) *, то на основании рис. 23-2,а легко записать:

$$I_y + I_p + I_n = \frac{U_1 - U_2}{R_0}, \quad (23-5a)$$

откуда

$$\Delta I_y + \Delta I_p + \Delta I_n = \frac{\Delta U_1 - \Delta U_2}{R_0}, \quad (23-5б)$$

Для энергетических расчетов, не связанных с дифференциальными параметрами, можно положить $U_2 = \text{const}$; $\Delta U_2 = 0$, т. е. считать стабилизатор идеальным. Кроме того, пренебрежем для простоты током I_y , поскольку в реальных стабилизаторах $I_y \ll I_p$ (в случае необходимости можно в дальнейших формулах понимать под I_p сумму $I_p + I_y$). Тогда из выражений (23-5) следует, что ток регулирующего элемента зависит от двух величин — тока нагрузки и входного напряжения. А именно, если $U_1 = \text{const}$, то изменение тока I_n сопровождается таким же, но противоположным по знаку изменением тока I_p ; иначе говоря, токи I_n и I_p меняются во взаимно противоположных направлениях:

$$\Delta I_p = -\Delta I_n. \quad (23-6a)$$

Если же, наоборот, $I_n = \text{const}$, то изменение тока I_p прямо пропорционально приращению входного напряжения:

$$\Delta I_p = \frac{\Delta U_1}{R_0}. \quad (23-6б)$$

Выражения (23-6) приводят к выводу, что ток регулирующего элемента минимален при максимальном токе нагрузки и минималь-

* Этот ток, как правило, несуществен. При необходимости можно считать сопротивление R_i подключенным параллельно сопротивлению R_n в виде неизменной нагрузки.

ном входном напряжении. Обозначим через $I_{p, \min}$ минимально допустимый ток регулирующего элемента¹. Тогда с учетом сказанного номинальный ток будет связан с номинальным входным напряжением следующим образом:

$$I_{p, \text{ном}} = I_{p, \min} + \frac{U_{1 \text{ном}} - U_{1 \min}}{R_0} = I_{p, \min} + \frac{\delta_{н1} U_{1 \text{ном}}}{R_0}. \quad (23-7)$$

Если сопротивление R_0 мало, ток $I_{p, \text{ном}}$ может намного превышать ток $I_{p, \min}$.

Из тех же выражений (23-6) ясно, что максимальный ток (по которому подбирают регулирующий элемент) имеет место при холостом ходе стабилизатора и максимальном входном напряжении. Прибавляя к току $I_{p, \min}$ приращения (23-6), получаем:

$$I_{p, \max} = I_{p, \min} + \frac{(\Delta U_1)_{\max}}{R_0} + I_{н, \max} = I_{1 \min} + \frac{U_{1 \text{ном}}}{R_0} \delta_1. \quad (23-8)$$

Здесь $I_{1 \min} = I_{p, \min} + I_{н, \max}$ — минимальный входной ток — величина, характерная для параллельных стабилизаторов, которая будет часто встречаться ниже.

Напряжение на регулирующем элементе в простейшем случае (рис. 23-2, а) равно выходному напряжению; следовательно, значение U_2 ограничено допустимым напряжением регулирующего элемента:

$$U_2 \leq U_{p, \text{доп}}. \quad (23-9)$$

Иногда последовательно с регулирующим элементом включают «гасящее» сопротивление R , назначение которого рассматривается ниже. В этом случае $U_p = U_2 - I_p R < U_2$. Однако наличие «гасящего» сопротивления не позволяет избежать ограничения (23-9), так как при случайных перегрузках, когда ток I_p уменьшается, напряжение U_p растет и регулирующий транзистор может пробиться.

Необходимое значение номинального входного напряжения легко найти из формулы (23-3а), если подставить $U_{1 \min} = U_2 + I_{1 \min} R_0$. Тогда

$$U_{1 \text{ном}} = \frac{1}{1 - \delta_{н1}} (U_2 + I_{1 \min} R_0). \quad (23-10)$$

Коэффициент полезного действия стабилизатора найдем из определения (23-2), используя полученное значение $U_{1 \text{ном}}$ и номинальный входной ток $I_{1 \text{ном}} = I_{н, \max} + I_{p, \text{ном}}$. Тогда с учетом (23-7)

$$\eta = \frac{(1 - \delta_{н1})^2 U_2 I_{н, \max}}{(U_2 + I_{1 \min} R_0) (I_{1 \min} + \delta_{н1} U_2 / R_0)}. \quad (23-11)$$

Легко убедиться, что знаменатель (23-11) имеет минимум, а к. п. д. — максимум при некотором оптимальном зна-

¹ Очевидно, что регулирующий элемент (транзистор) не может работать при сколь угодно малых токах из-за уменьшения коэффициента усиления, роста выходного сопротивления и т. п. Ток $I_{p, \min}$ выбирается из условий сохранения приемлемых параметров регулирующего элемента.

чении R_0

$$R_{0\text{ опт}} = \frac{V\delta_{н1}U_2}{I_{1\text{ мин}}} \quad (23-12)$$

Оптимальному сопротивлению (которое чаще всего принимают в сколько-нибудь мощных стабилизаторах) соответствуют максимальный к. п. д.

$$\eta_{\text{макс}} = \frac{(1 - \sqrt{\delta_{н1}})^2 I_{н. \text{ макс}}}{I_{1\text{ мин}}} \approx (1 - \sqrt{\delta_{н1}})^2 \quad (23-13)$$

и номинальные параметры в оптимальном режиме

$$U_{1\text{ ном}} = \frac{U_2}{1 - \sqrt{\delta_{н1}}}; \quad (23-14)$$

$$I_{р. \text{ ном}} = I_{р. \text{ мин}} + \frac{V\delta_{н1}}{1 - \sqrt{\delta_{н1}}} I_{1\text{ мин}}. \quad (23-15)$$

Функции допуска $\delta_{н1}$, необходимые для расчетов, показаны на рис. 23-3.

Может показаться, что уменьшение допуска $\delta_{н1}$ выгодно с точки зрения к. п. д. и других параметров. На самом деле это не так: с уменьшением $\delta_{н1}$ уменьшается $R_{0\text{ опт}}$ и соответственно возрастает ток $I_{р. \text{ макс}}$. Это, как показано ниже, заставляет выбирать регулирующий элемент с большим запасом по току и мощности. При обычных значениях $\delta_{н1} = 0,1 \div 0,2$ получаются довольно низкие значения к. п. д. $\eta_{\text{макс}} \approx 0,5 \div 0,3$. При токах $I_{н} < I_{н. \text{ макс}}$ оптимальный режим нарушается и к. п. д. становится еще меньше.

Выбор регулирующего элемента для стабилизатора основывается прежде всего на том, чтобы его допустимый ток превышал $I_{р. \text{ макс}}$. Подставляя (23-10) в (23-8) и полагая $I_{р. \text{ доп}} \geq I_{р. \text{ макс}}$, получаем:

$$I_{р. \text{ доп}} \geq \frac{1 + \delta_{в1}}{1 - \delta_{н1}} I_{1\text{ мин}} + \frac{\delta_{л}}{1 - \delta_{н1}} \frac{U_2}{R_0}. \quad (23-16a)$$

При оптимальном режиме, используя (23-12), получаем:

$$I_{р. \text{ доп}} \geq \frac{1 + \delta_{в1}/\sqrt{\delta_{н1}}}{1 - \sqrt{\delta_{н1}}} I_{1\text{ мин}}. \quad (23-16б)$$

Функции допусков $\delta_{в1}$ и $\delta_{н1}$, необходимые для расчетов, показаны на рис. 23-3 пунктиром.

С увеличением R_0 допустимый ток согласно (23-16a) уменьшается; при $R_0 \rightarrow \infty$ он на 30—35% меньше, чем в оптимальном режиме. Значения $R_0 < R_{0\text{ опт}}$ не следует применять на практике, так как они, не давая никаких преимуществ, уменьшают коэффициент стабилизации (см. ниже).

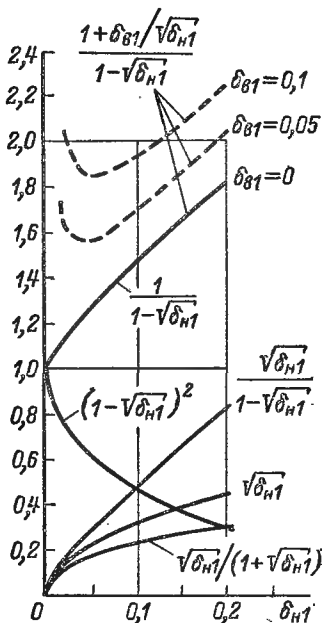


Рис. 23-3. Функции допусков δ , используемые в расчетных формулах.

Мощность регулирующего элемента оценивается по тем же формулам (23-16), если заменить $I_{p. \text{доп}}$ на $P_{p. \text{доп}}$, а правую часть умножить на U_2 .

Необходимую мощность $P_{p. \text{доп}}$ можно существенно уменьшить, если последовательно с регулирующим элементом включить «гасящее» сопротивление (см. приложение 1 в [176]). Тогда увеличение тока I_p будет сопровождаться уменьшением напряжения U_p . Максимальная рассеиваемая мощность $P_{p. \text{макс}}$ получается при $U_p = U_2/2$ и составляет $U_2^2/4R$. Сопротивление R выбирают из условия $I_{p. \text{макс}} R \leq U_2$, чтобы транзистор не вошел в насыщение даже при максимальном токе. Положим $R = U_2/I_{p. \text{макс}}$; тогда мощность $P_{p. \text{макс}} = 1/4 U_2 I_{p. \text{макс}}$, т. е. в 4 раза меньше, чем без «гасящего» сопротивления. Соответственно при одном и том же максимальном токе $I_{p. \text{макс}}$ можно выбирать менее мощный транзистор.

Перейдем к рассмотрению дифференциальных параметров стабилизатора — выходного сопротивления и коэффициента стабилизации.

Выходное сопротивление найдем из общего определения (23-1а). Для этого из соотношения (23-5б) выразим приращение ΔI_n через ΔU_2 , полагая $\Delta U_1 = 0$ и подставляя

$$\Delta I_y = \Delta U_2 / R_y; \quad \Delta I_p = K_i \Delta I_y$$

(R_y — сумма входного сопротивления усилительного элемента и сопротивления опорного элемента; K_i — суммарный коэффициент усиления тока в усилительном и регулирующем элементах). После элементарных преобразований получаем:

$$R_{\text{вых}} = R_0 \parallel \left(\frac{R_y}{K_i + 1} \right).$$

Эквивалентное сопротивление в круглых скобках играет важную роль в теории стабилизаторов. Назовем его *характеристическим сопротивлением* и обозначим символом

$$R_\infty = \frac{R_y}{K_i + 1}. \quad (23-17)$$

По физическому смыслу — это выходное сопротивление активной части стабилизатора (т. е. $R_{\text{вых}}$ при $R_0 \parallel R_i = \infty$).

Ниже мы убедимся, что в реальных случаях всегда соблюдается условие $R_0 \gg R_\infty$, поэтому

$$R_{\text{вых}} = R_0 \parallel R_\infty \approx R_\infty. \quad (23-18)$$

Коэффициент стабилизации найдем из общего определения (23-1б). Для этого из соотношения (23-5б) выразим отношение $\Delta U_1 / \Delta U_2$, подставив приведенные выше значения ΔI_y и ΔI_p , а также $\Delta I_n = \Delta U_2 / R_n$. После преобразований получим следующее общее выражение:

$$K = \frac{U_2}{U_1} \frac{R_0}{R_0 \parallel R_n \parallel R_\infty}. \quad (23-19а)$$

Обычно выполняются условия

$$R_0 \gg R_{\infty}; \quad R_H \gg R_{\infty},$$

первое из которых использовалось выше [см. (23-18)]. Тогда получается упрощенное выражение:

$$K = \frac{U_2}{U_1} \frac{R_0}{R_{\infty}}. \quad (23-196)$$

Наконец, подставляя в (23-196) значение U_1 из (23-10), запишем коэффициент стабилизации следующим образом:

$$K = \frac{(1 - \delta_{H1}) U_2 R_0}{(U_2 + I_{1\text{мин}} R_0) R_{\infty}}. \quad (23-19в)$$

Как видно, с увеличением сопротивления R_0 значение K возрастает и при $R_0 \rightarrow \infty$ стремится к предельному значению:

$$K_{\infty} = \frac{(1 - \delta_{H1}) U_2}{I_{1\text{мин}} R_{\infty}}. \quad (23-20)$$

При оптимальном балластном сопротивлении (23-12) выражение (23-19в) дает:

$$K_{\text{опт}} = \frac{\sqrt{\delta_{H1}} (1 - \sqrt{\delta_{H1}}) U_2}{I_{1\text{мин}} R_{\infty}} = \frac{\sqrt{\delta_{H1}}}{1 + \sqrt{\delta_{H1}}} K_{\infty}. \quad (23-21)$$

Следовательно, в обоих случаях — и в предельном, и в оптимальном режимах — коэффициент стабилизации в первую очередь зависит от характеристического сопротивления. Последнее, естественно, нужно делать достаточно малым (это же желательно и с точки зрения выходного сопротивления).

Уменьшение сопротивления R_{∞} согласно (23-17) достигается прежде всего увеличением коэффициента усиления по току. Заметим также, что коэффициент стабилизации возрастает с уменьшением тока нагрузки — главной составляющей тока $I_{1\text{мин}}$.

В заключение рассмотрим более общую схему стабилизатора, в которой помимо обычной обратной связи через элемент 2 имеется прямая связь между входом и регулирующим элементом через элемент 4 (рис. 23-4). В такой схеме приращение регулирующего тока имеет вид:

$$\Delta I_p = (\Delta U_2 K_{i2}) / R_{y2} + (\Delta U_1 K_{i4}) / R_{y4},$$

где индексы 2 и 4 относятся соответственно к цепям обратной и прямой связи. Тогда из соотношения (23-56) нетрудно получить величину $\Delta U_1 / \Delta U_2$ и далее — коэффициент стабилизации:

$$K = \frac{U_2}{U_1} \frac{R_0 / R_{\infty}}{1 - K_{i4} R_0 / R_{y4}}. \quad (23-22)$$

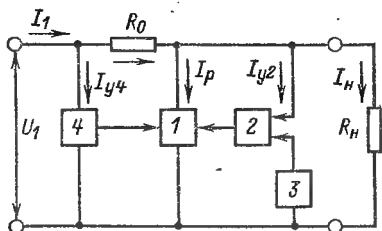


Рис. 23-4. Обобщенная скелетная схема параллельного стабилизатора с цепью прямой связи.

1 — регулирующий элемент; 2 — сравнивающий и усилительный элемент цепи обратной связи; 3 — опорный элемент; 4 — элемент прямой связи.

Сравнивая (23-22) с выражением (23-19б), можно сделать следующие выводы. Если $K_{I_d} < 0$, т. е. если прямая связь противодействует обратной, то коэффициент стабилизации будет меньше, чем при отсутствии прямой связи. Если же $K_{I_d} > 0$ (случай, показанный на рис. 23-4), то коэффициент стабилизации повысится. Делая второй член в знаменателе (23-22) близким к единице, можно получить любые значения K . Однако при этом малейшие изменения величин K_{I_d} и R_{y_d} будут весьма сильно менять значение K . Такой режим нежелателен.

Величина выходного сопротивления практически не зависит от наличия элемента прямой связи и выражается формулой (23-18).

Диодные стабилизаторы. Полупроводниковые стабилитроны (см. § 3-3) позволяют осуществлять простейшие схемы стабилизации без использования усилительных приборов. Схема стабилизатора на одном стабилитроне показана на рис. 23-5. Очевидно, что эта схема соответствует структуре параллельных стабилизаторов, хотя регулирующий и усилительный элементы отсутствуют, а стабилитрон D совмещает функции опорного и регулирующего элементов. Анализ и расчет такого стабилизатора можно вести по

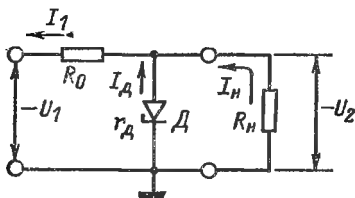


Рис. 23-5. Диодный стабилизатор.

формулам предыдущего раздела, полагая $K_i = 0$; $R_y = r_d$ (где r_d — дифференциальное сопротивление стабилитрона) и заменяя I_p на I_d .*

Стабилитрон подбирают по значению выходного напряжения ($U_d = U_2$), а также, если возможно, по значению тока нагрузки из неравенств (23-16). После этого становится известным параметр r_d . Согласно (23-17) при $K_i = 0$ имеем:

$$R_{\infty} = r_d,$$

т. е. в диодных стабилизаторах характеристическое сопротивление по существу является величиной заданной. Соответственно заданными оказываются и важнейшие параметры стабилизаторов: выходное сопротивление (23-18)

$$R_{\text{вых}} \approx r_d$$

и предельный коэффициент стабилизации (23-20)

$$K_{\infty} = \frac{(1 - \delta_{H1}) U_2}{(I_{d, \text{мин}} + I_{H, \text{макс}}) r_d}.$$

Ток $I_{d, \text{мин}}$ обычно полагают равным 2—3 мА для маломощных стабилитронов и 3—5 мА для мощных. Сопротивление r_d , которое, вообще говоря, зависит от тока, обычно считают постоянным, равным номинальному значению.

* Последнее объясняется тем, что при общем анализе мы для простоты писали I_p вместо суммы $I_p + I_y$. Однако в диодном стабилизаторе $I_p = 0$. Следовательно, вместо суммы $I_y + I_p$ нужно использовать ток $I_y = I_d$.

Все сказанное свидетельствует о том, что диодный стабилизатор не обладает должной гибкостью и с его помощью трудно, а часто невозможно удовлетворить многим техническим заданиям, в которых комбинации величин U_2 , I_n , $R_{\text{вых}}$, K могут быть самыми различными.

Простой способ повысить коэффициент стабилизации при прочих равных условиях состоит в использовании каскадных схем (рис. 23-6). Идея таких схем очевидна, поскольку входное напряжение второго каскада стабилизировано первым каскадом [177]. Расчет каскадного стабилизатора осуществляется «справа налево», т. е. от выхода к входу. Легко убедиться, что общий коэффициент стабилизации двухкаскадной схемы равен:

$$K = \frac{(1 - \delta_{n1}) U_2 R_{01} R_{02}}{[U_2 + I_{n, \text{макс}} (R_{01} + R_{02}) + I_{d, \text{мин}} (2R_{01} + R_{02})] r_{d1} r_{d2}}, \quad (23-23)$$

где для второго каскада положено $\delta_{n1} = 0$. Номинальное напряжение стабилитрона D_1 составляет:

$$U_{d1} = U_2 + (I_{n, \text{макс}} + I_{d, \text{мин}}) R_{02}. \quad (23-24)$$

Например, если в качестве стабилитрона D_2 принять стабилитрон типа Д814В, то при $R_{02} = 175 \text{ Ом}$ и $I_{n, \text{макс}} = 16 \text{ мА}$ получим $U_{d1} = 13,5 \text{ В}$. В качестве D_1 можно принять стабилитрон типа Д814Д, имеющий $r_d = 15 \text{ Ом}$. Тогда при $\delta_{n1} = 0,1$ и $R_{01, \text{опт}} \approx 200 \text{ Ом}$ находим $K \approx 120$ (на порядок больше, чем в простейшей однокаскадной схеме на стабилитроне Д814В). Выходное сопротивление определяется вторым каскадом и остается равным $r_{d2} = 10 \text{ Ом}$.

Диодные стабилизаторы используются главным образом в качестве источников опорного напряжения для более мощных транзисторных стабилизаторов (рис. 23-7, 23-9 и др.) или в качестве источника питания для слаботочных схем, например цепей смещения. В этих случаях иногда удается обеспечить условие $I_{n, \text{макс}} \leq I_{d, \text{мин}}$, при котором стабильность по напряжению U_1 и току I_n может быть весьма высокой.

При $I_{n, \text{макс}} \rightarrow 0$, т. е. на холостом ходу, предельный коэффициент стабилизации согласно (23-20) имеет максимальное значение

$$K_{\infty \text{ макс}} = \frac{(1 - \delta_{n1}) U_2}{I_{d, \text{мин}} r_d}. \quad (23-25a)$$

Наоборот, при полной нагрузке, когда $I_{1, \text{мин}} \approx I_{n, \text{макс}} \approx 0,5 I_{d, \text{доп}}$ [см. выражения (23-16) и рис. 23-3], коэффициент стабилизации согласно (23-20) имеет минимальное значение

$$K_{\infty \text{ мин}} \approx \frac{2(1 - \delta_{n1}) U_2}{I_{d, \text{доп}} r_d}. \quad (23-256)$$

Расчеты по формулам (23-25) показывают, что минимальный коэффициент стабилизации не превышает 50 даже у слаботочных

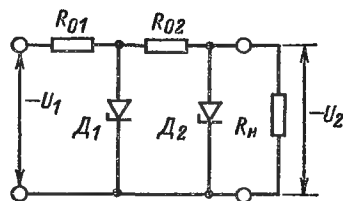


Рис. 23-6. Двухкаскадный диодный стабилизатор.

стабилизаторов, а у сильноточных он составляет всего 15—40. Максимальный коэффициент стабилизации у слаботочных стабилизаторов лежит в пределах 150—750, а у сильноточных достигает 1000—5000.

Температурный и временной дрейф диодных стабилизаторов такой же, как у отдельного стабилитрона (см. § 3-3). В широком диапазоне температур дрейф выходного напряжения может составлять до 10% и больше, что намного превышает нестабильность по входному напряжению и току нагрузки.

В последнее время, особенно в микроэлектронике, получили распространение диодные стабилизаторы, у которых вместо стабилитрона используется кремниевый $p-n$ переход, включенный в прямом направлении (см. рис. 2-21,б). В частности, это может быть эмиттерный переход транзистора (при этом коллектор обычно соединяют с базой). Такие стабилизаторы на холостом ходу обеспечивают коэффициент стабилизации, равный 15—20 при напряжениях 0,6—0,7 В; выходное сопротивление определяется рабочим током согласно (2-38).

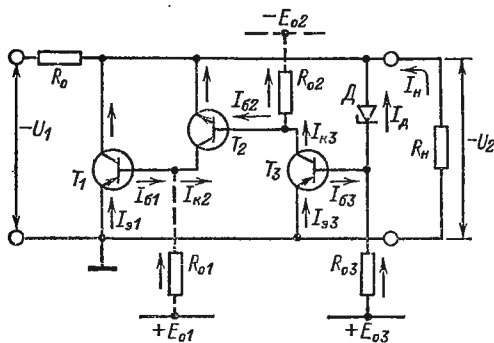


Рис. 23-7. Трехкаскадный параллельный стабилизатор.

Анализ показывает, что однокаскадный параллельный стабилизатор не имеет преимуществ перед диодным, а двухкаскадный уступает по параметрам двухкаскадному последовательному. Поэтому рассмотрим только трехкаскадную схему (рис. 23-7).

Сравнивая эту схему со скелетной схемой на рис. 23-2,а, замечаем, что в данном случае опорный элемент представлен стабилитроном D , сравнивающий элемент — усилителем тока на транзисторах T_3 , T_2 и регулирующий элемент — мощным транзистором T_1 . Соответственно

$$\left. \begin{aligned} I_p &= I_{к1}; \\ R_y &= r_d + R_{вх3}; \\ K_i &= \beta_1 \beta_2 \beta_3, \end{aligned} \right\} \quad (23-26)$$

где $R_{вх3} = r_{б3} + (1 + \beta_3)r_{э3}$.

Тогда согласно (23-17) характеристическое сопротивление, а вместе с ним и выходное сопротивление стабилизатора будут равны:

$$R_{вых} = R_{\infty} = \frac{r_d + r_{б3} + (1 + \beta_3)r_{э3}}{1 + \beta_1 \beta_2 \beta_3}. \quad (23-27)$$

Коэффициент стабилизации в оптимальном режиме, свойственном мощным стабилизаторам, согласно (23-21) равен:

$$K_{\text{опт}} = \frac{V \delta_{\text{н1}} (1 - V \delta_{\text{н1}}) U_2}{(I_{\text{н. макс}} + I_{\text{к1 макс}}) R_{\infty}}. \quad (23-28)$$

Выходное напряжение практически равно напряжению стабилизатора:

$$U_2 = U_{\text{д}} + U_{\text{бз}} \approx U_{\text{д}}, \quad (23-29)$$

так как во всех реальных схемах $U_{\text{бз}} \ll U_{\text{д}}$.

Токоотводящие резисторы R_0 вместе с источниками питания E_0 обеспечивают минимальный остаточный ток регулирующего транзистора и позволяют снизить номинальную мощность усилительных транзисторов.

В самом деле, в отсутствие резистора R_0 ток транзистора T_1 (при максимальной нагрузке) составлял бы:

$$I_{\text{к1}} \approx \beta_1 \beta_2 \beta_3 I_{\text{д. мин.}}$$

При $I_{\text{д. мин}} = 2 \text{ мА}$ и $\beta_{1,2,3} = 30$ получается $I_{\text{к1}} \approx 54 \text{ А}$, что, конечно, неприемлемо. При наличии резистора R_{01} , сопротивление которого выбрано из условия¹

$$I_{01} = \frac{E_{01}}{R_{01}} = I_{\text{к2}} - I_{\text{б1}}, \quad (23-30)$$

транзистор T_1 находится в нормальном режиме.

Полагая $I_{\text{эз}} = 1 \text{ мА}$, можно считать $r_{\text{эз}} = 25 \text{ Ом}$; тогда при $r_{\text{д}} = 10 \text{ Ом}$; $r_{\text{бз}} = 100 \text{ Ом}$ и $\beta_{1,2,3} = 30$ получим из (22-26):

$$R_{\text{вых}} \approx 0,028 + 0,004 = 0,032 \text{ Ом}.$$

Как видим, определяющую роль в значении $R_{\text{вых}}$ играет слагаемое, содержащее $r_{\text{эз}}$. Для того чтобы оба слагаемых были равноценными, нужно увеличить ток $I_{\text{эз}}$ примерно на порядок, т. е. до максимального значения, свойственного маломощным транзисторам. В этом случае можно получить $R_{\text{вых}} < 0,01 \text{ Ом}$, т. е. на 2—3 порядка меньше, чем в диодном стабилизаторе.

При использовании мощных транзисторов ток нагрузки может достигать до нескольких ампер, а при параллельном соединении транзисторов — до 10 А и выше. Для таких токов балластное сопротивление может иметь столь малое значение (1—2 Ом), что его роль вполне может играть выходное сопротивление выпрямителя. Более того, последнее нередко превышает необходимое значение R_0 и тогда стабилизатор не может работать в оптимальном режиме, если не пересчитать выпрямитель.

¹ В этом условии считается, что напряжение $U_{\text{бз}} = 0$, поэтому

$$I_{01} = \frac{E_{01} - U_{\text{б1}}}{R_{01}} \approx \frac{E_{01}}{R_{01}}.$$

Базовым и коллекторным токами в правой части (23-30) задаются, исходя из желательного режима транзисторов.

Сопротивления R_{02} и R_{03} выбирают из аналогичных условий.

Общие свойства. Из рис. 23-2, б легко записать соотношение

$$I_y + I_n = I_p + \frac{U_1 - U_2}{R_i}, \quad (23-31a)$$

откуда следует, что

$$\Delta I_y + \Delta I_n - \Delta I_p = \frac{\Delta U_1 - \Delta U_2}{R_i}. \quad (23-31б)$$

Для энергетических расчетов, не связанных с дифференциальными параметрами, можно, как и в параллельных стабилизаторах, положить $U_2 = \text{const}$; $\Delta U_2 = 0$, а также для простоты пренебречь током I_y (в случае необходимости следует в дальнейших выражениях понимать под I_p сумму $I_p + I_y$). Внутреннее сопротивление регулирующего элемента R_i играет в последовательных стабилизаторах принципиальную роль и пренебрегать им, как это делалось в предыдущем параграфе, нельзя. Это ясно хотя бы из того, что сопротивление R_i занимает в формулах (23-31) то же место, что и весьма важное балластное сопротивление R_0 в формулах (23-5).

Из выражений (23-31) следует, что в последовательных стабилизаторах при $U_1 = \text{const}$ ток регулирующего элемента повторяет изменения тока нагрузки:

$$\Delta I_p = \Delta I_n, \quad (23-32a)$$

а при $I_n = \text{const}$ изменения тока I_p равны, но противоположны по знаку изменениям тока в сопротивлении R_i :

$$\Delta I_p = - \frac{\Delta U_1}{R_i}. \quad (23-32б)$$

Выражения (23-32) сходны с выражениями (23-6) для параллельных стабилизаторов, но разница в знаках приводит к совершенно иным выводам. Во-первых, в данном случае максимальный ток $I_{p, \text{макс}}$, по которому подбирают регулирующий элемент, имеет место при максимальном токе нагрузки и минимальном входном напряжении:

$$I_{p, \text{макс}} = I_{n, \text{макс}} - \frac{U_{1\text{мин}} - U_2}{R_i} = I_{n, \text{макс}} - \frac{(1 - \delta_{н1}) U_{1\text{ном}} - U_2}{R_i} \quad (23-33)$$

(часто с некоторым запасом считают $I_{p, \text{макс}} = I_{n, \text{макс}}$, поскольку сопротивление R_i сравнительно велико). Во-вторых, *последовательные стабилизаторы не могут работать при холостом ходе*: при $I_n = 0$ из (23-31a) получается $I_p < 0$. В-третьих, минимальный, «остаточный» ток $I_{p, \text{мин}}$, необходимый для нормальной работы регулирующего элемента, получается при минимальном токе нагрузки и максимальном входном напряжении, т. е. согласно (23-31a)

$$I_{p, \text{мин}} = I_{n, \text{мин}} - \frac{U_{1\text{макс}} - U_2}{R_i} = I_{n, \text{мин}} - \frac{(1 + \delta_{н1}) U_{1\text{ном}} - U_2}{R_i}. \quad (23-34)$$

Ток $I_{н. мин}$ можно беспроблемно, подключив на выходе стабилизатора постоянное сопротивление — шунт. Тогда по отношению к $U_{н. макс}$ в формуле (23-33) следует понимать сумму тока в собственно нагрузке и тока в шунте ($I_{ш} = I_{н. мин}$).

В рабочем режиме напряжение на регулирующем элементе составляет $U_p = U_1 - U_2$. Однако в момент включения стабилизатора, учитывая наличие конденсатора на выходе (см. с. 647), получается $U_p = U_1$; такой же результат будет при кратковременном коротком замыкании выходных зажимов. Поэтому при выборе регулирующего элемента по напряжению ориентируются на условие $U_{p. макс} = U_{1 макс}$.

Сначала запишем очевидное соотношение

$$U_{1 мин} = U_2 + U_{p. мин}, \quad (23-35)$$

где $U_{p. мин}$ — напряжение, при котором регулирующий элемент еще работоспособен¹. Затем, используя формулы (23-3а) и (23-3б), легко найдем напряжение $U_{1 макс}$ и равное ему напряжение $U_{p. макс}$:

$$U_{p макс} = U_{1 макс} = \frac{1 + \delta_{в1}}{1 - \delta_{н1}} (U_2 + U_{p. мин}). \quad (23-36)$$

Регулирующий элемент выбирается из условия $U_{p. доп} \geq U_{p. макс}$. Если, наоборот, задаться типом регулирующего элемента, т. е. значением $U_{p. доп}$, то из условия $U_{p. макс} \leq U_{p. доп}$ получаем ограничение на значение U_2 :

$$U_2 \leq \frac{1 - \delta_{н1}}{1 + \delta_{в1}} U_{p. доп} - U_{p. мин}. \quad (23-37)$$

Например, если $U_{p. доп} = 30$ В; $U_{p. мин} = 1$ В и $\delta_{н1} = \delta_{в1} = 0,1$, то $U_2 \leq 23,5$ В.

Если при выборе $U_{p. доп}$ исходить из рабочего режима стабилизатора, т. е. из условия $U_{p. макс} = U_{1 макс} - U_2$, то необходимое значение $U_{p. доп}$ оказывается значительно ниже, а при заданном регулирующем элементе можно реализовать гораздо большие значения U_2 . Например, при тех же исходных параметрах, что и в предыдущем примере, получаем $U_2 \leq 130$ В.

Значение номинального входного напряжения легко определяется из соотношения (23-35) с учетом (23-3а):

$$U_{1 ном} = \frac{U_2 + U_{p. мин}}{1 - \delta_{н1}}. \quad (23-38)$$

Коэффициент полезного действия можно найти из определения (23-2), учитывая, что ток $I_{1 ном}$ складывается из максимального тока истинной нагрузки и тока в шунте: $I_{1 ном} = I_{н. макс} + I_{ш}$.

¹ Значение $U_{p. мин}$ должно превышать сумму напряжений насыщения, пульсаций, наводок и толчков на входе стабилизатора. При наличии хорошего емкостного фильтра в выпрямителе обычно можно считать $U_{p. мин} = 1$ В.

Подставляя $I_{ш} = I_{н. мин}$ из (23-34) в (23-2) и используя (23-38), после небольших преобразований получаем:

$$\eta = (1 - \delta_{н1}) \frac{U_2}{U_2 + U_{р. мин}} \frac{I_{н. макс}}{I_{н. макс} + I_{р. мин} + \frac{\delta_1 U_2 + (1 + \delta_{в1}) U_{р. мин}}{(1 - \delta_{н1}) R_i}}. \quad (23-39)$$

Оптимальный режим в данном случае не имеет места, так как напряжение U_1 не зависит от R_i и к. п. д. монотонно возрастает с увеличением R_i . В пределе, при $U_{р. мин} = 0$; $I_{р. мин} = 0$ и $R_i = \infty$, имеем $\eta_{макс} = (1 - \delta_{н1})$. Если $\delta_{н1} = 0,1 \div 0,2$, то $\eta_{макс} = 0,9 \div 0,8$. Практически $\eta = 0,6 \div 0,8$. С уменьшением тока нагрузки к. п. д. уменьшается, но эта зависимость проявляется слабо до тех пор, пока ток $I_{н. макс}$ не делается сравнимым с двумя другими слагаемыми в знаменателе (23-39).

Выбор регулирующего элемента по току осуществляется из условия $I_{р. доп} \geq I_{р. макс}$. Подставляя (23-33) с учетом (23-38), получаем:

$$I_{р. доп} \geq I_{н. макс} - \frac{U_{р. мин}}{R_i}. \quad (23-40)$$

Практически считают $I_{р. доп} \geq I_{н. макс}$. Мощность регулирующего элемента оценивается, исходя из величины $I_{р. макс} U_{р. макс}$, где $U_{р. макс} = U_{1 макс} - U_2$. Используя (23-33) и (23-36), приходим к условию

$$P_{р. доп} \geq \left(I_{н. макс} - \frac{U_{р. мин}}{R_i} \right) \frac{\delta_1 U_2 + (1 + \delta_{в1}) U_{р. мин}}{1 - \delta_{н1}}. \quad (23-41)$$

Если $U_{р. мин} = 0$ и $\delta_{в1} = \delta_{н1} = 0,1$, то $P_{р. доп} \geq 0,22 U_2 I_{н. макс}$.

Перейдем к дифференциальным параметрам. Выходное сопротивление можно найти из общего определения (23-1а) и соотношения (23-31б), полагая $\Delta U_1 = 0$; $\Delta I_y = \Delta U_2 / R_y$ и $\Delta I_p = -K_i I_y^*$. После подстановок получим:

$$R_{вых} = R_i \parallel R_{\infty} \approx R_{\infty}, \quad (23-42)$$

где R_{∞} — характеристическое сопротивление (23-17). Сравнивая (23-42) и (23-18), видим, что различие состоит лишь в замене R_0 на R_i . Это можно было предсказать заранее, поскольку при закороченных входных зажимах обе схемы на рис. 23-2 имеют одинаковую структуру.

Коэффициент стабилизации легко найти из определения (23-1б), если выразить величину $\Delta U_1 / \Delta U_2$ с помощью (23-31б) при указанных выше значениях ΔI_y и ΔI_p :

$$K = \frac{U_2}{U_1} \frac{R_i}{R_i \parallel R_{н1} \parallel R_{\infty}}. \quad (23-43а)$$

* Знак «—» у коэффициента усиления отражает тот факт, что в последовательных стабилизаторах в отличие от параллельных направления токов I_y и I_p относительно выходного зажима противоположны друг другу (см. рис. 23-2).

Сравнивая (23-43а) с (23-19а), снова замечаем, что различие состоит в замене R_0 на R_i .

В последовательных стабилизаторах, так же как и в параллельных, выполняются неравенства $R_n \gg R_\infty$ и $R_i \gg R_\infty$; поэтому практически коэффициент стабилизации имеет вид:

$$K = \frac{U_2}{U_1} \frac{R_i}{R_\infty}. \quad (23-43б)$$

Выражения (23-43) подчеркивают ту принципиальную роль, которую играет сопротивление R_i в последовательных стабилизаторах. Именно оно обуславливает конечное значение K , так как при $R_i = \infty$ вход и выход стабилизатора были бы изолированы друг от друга, входное напряжение никак не влияло бы на выходное и мы имели бы $K = \infty$. Сопротивление R_i , будучи параметром регулирующего элемента, оказывается заданной величиной после того, как регулирующий элемент выбран.

Подставляя (23-38) в (23-43б), получаем:

$$K = \frac{(1 - \delta_{н1}) U_2 R_i}{(U_2 + U_{p, \min}) R_\infty}. \quad (23-43в)$$

Поскольку $U_{p, \min}$ не зависит от R_i и поскольку R_i — величина заданная, понятие предельного коэффициента стабилизации в последовательных схемах отсутствует. Кроме того, коэффициент K , как видим, не зависит от тока нагрузки. Следовательно, единственным путем повышения стабильности по входному напряжению является уменьшение характеристического сопротивления.

В заключение остановимся на более общей схеме последовательного стабилизатора (рис. 23-8), в которой, как и на рис. 23-4, имеется не только обратная, но и прямая связь¹. Отмечая параметры цепи обратной связи индексом 2, а параметры цепи прямой связи индексом 4, легко показать, что выходное сопротивление (23-42) сохраняет прежнюю величину, а коэффициент стабилизации (23-43) следует разделить на двучлен:

$$1 + K_{14} R_i / R_{y4}, \quad (23-44)$$

знакомый по выражению (23-22)). По причинам, изложенным в § 23-2, прямую связь обычно стараются исключить, хотя при $K_{14} < 0$ она может повысить коэффициент стабилизации.

Однокаскадный стабилизатор. Стабилизатор, показанный на рис. 23-9, состоит из усилительной части, представленной тран-

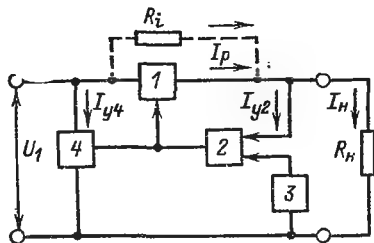


Рис. 23-8. Обобщенная скелетная схема последовательного стабилизатора с цепью прямой связи.

1 — регулирующий элемент; 2 — сравнивающий и усилительный элемент цепи обратной связи; 3 — опорный элемент; 4 — элемент прямой связи.

¹ Этот случай впервые проанализирован в работе [177], но результаты представлены в ней в несколько иной форме.

зистором T , и опорной части, представленной стабилитроном D , балластным сопротивлением R_0 и стабилизированным напряжением E_0 . Выходное напряжение стабилизатора равно опорному с точностью до малой величины $U_{эб}$. В сущности, эта схема является эмиттерным повторителем, у которого потенциал базы стабилизирован стабилитроном и равен U_d , а напряжение коллекторного питания U_1 меняется в определенных пределах. Поэтому выходное сопротивление стабилизатора можно было бы выразить с помощью формулы (9-6б), заменив R_r на r_d , но мы воспользуемся теми общими методами, которые описаны в предыдущем разделе.

Сравнивая эквивалентную схему на рис. 23-9, б с общей скелетной схемой на рис. 23-2, б, приходим к соотношениям $I_p = I_k$; $R_y = R_B$; $K_i = \beta$, где $R_B \approx r_6 + r_d$ (поскольку $R_0 \gg r_d$). Используя (23-17), найдем характеристическое сопротивление между

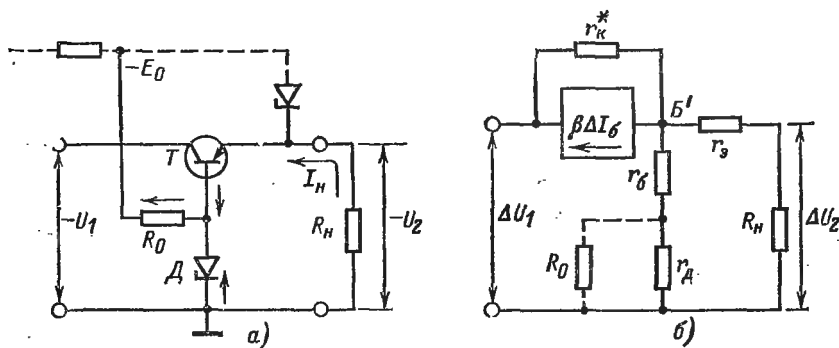


Рис. 23-9. Однокаскадный последовательный стабилизатор с питанием стабилитрона от стабилизированного напряжения.

а — принципиальная схема; б — эквивалентная схема.

точками $B' - O$ и, добавляя к нему величину $r_э$, получим полное характеристическое сопротивление, а вместе с ним и выходное сопротивление стабилизатора:

$$R_{\text{вых}} \approx R_{\infty} = r_э + \frac{r_d + r_6}{1 + \beta}. \quad (23-45)$$

Коэффициент стабилизации легко найти из формулы (23-43в), если подставить в нее найденное значение R_{∞} и $R_i = r_k^*$. Умножив числитель и знаменатель на $1 + \beta$ и учитывая, что $(1 + \beta) r_k^* = r_k$, получаем:

$$K \approx \frac{(1 - \delta_{н1}) U_2 r_k}{[r_d + r_6 + (1 + \beta) r_э] (U_2 + U_{\text{к. мин}})}. \quad (23-46)$$

Пусть, например, $r_k = 5 \text{ кОм}$, $r_6 = 20 \text{ Ом}$; $\beta = 30$ (параметры, характерные для транзисторов средней мощности) и пусть $I_э = 250 \text{ мА}$; $r_d = 10 \text{ Ом}$; $U_{\text{к. мин}} = 0,1 U_2$ и $\delta_{н1} = 0,1$; тогда $R_{\text{вых}} \approx 1 \text{ Ом}$ и $K = 125$. Последнее значение часто вполне приемлемо, но выходное сопротивление

ние сравнительно велико. Именно величина $R_{\text{вых}}$ ограничивает максимальный ток нагрузки в однокаскадном стабилизаторе, поскольку при больших значениях $I_{\text{н. макс}}$ получается большая неустойчивость по току δI_2 [см. (23-46)].

Энергетический расчет стабилизатора проводится по общим формулам предыдущего раздела. Балластный резистор R_0 стабилитрона D является одновременно токоотводящим: через него идет сумма токов I_6 и I_d , направленных навстречу друг другу. Значение R_0 выбирают из очевидного условия

$$R_0 = \frac{E_0 - U_6}{I_d + I_6}, \quad (23-47)$$

где $U_6 \approx U_2$; $I_6 \approx I_{\text{н. макс}}/\beta$, а током I_d можно задаться, соблюдая неравенство

$$I_{d, \text{ мин}} < I_d < I_{d, \text{ макс}} - I_6.$$

Правая часть этого неравенства учитывает тот факт, что при сбросе нагрузки ток базы уменьшается почти до нуля, а ток стабилитрона возрастает на величину I_y , поскольку $I_6 + I_d = \text{const}$. При этом величина $I_d + I_6$ не должна превышать $I_{d, \text{ макс}}$.

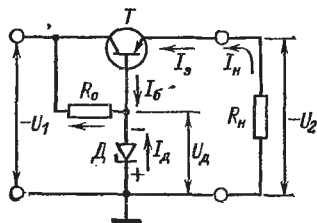


Рис. 23-10. Однокаскадный последовательный стабилизатор с питанием стабилитрона от входного напряжения.

Казалось бы, питание стабилитрона проще осуществить не от специального источника E_0 , а от источника входного напряжения (рис. 23-10). Такое решение возможно [177], но нецелесообразно, так как при этом резко снижается коэффициент стабилизации. В самом деле, если пренебречь изменениями напряжения $U_{эб}$, то стабильность потенциала эмиттера будет такой же, как стабильность потенциала базы. Значит, коэффициент стабилизации всей схемы будет близок к коэффициенту стабилизации ее опорной части, а у диодных стабилизаторов, как отмечалось в § 23-2, значение K не превышает 20—30 даже при больших балластных сопротивлениях. Между тем сопротивление R_0 в данном случае имеет меньшее значение, чем (23-47), так как оно должно удовлетворять условию

$$(I_6 + I_d) R_0 = U_1 - U_6 \approx U_1 - U_2,$$

где $U_1 < E_0$. Обычно значение R_0 на рис. 23-10 составляет 100—200 Ом, что при $r_d = 10$ Ом дает $K < 10 \div 20$.

Напряжение E_0 на рис. 23-9,а не является абсолютно постоянным, как считалось при составлении эквивалентной схемы (рис. 23-9, б), так как диодный стабилизатор, показанный пунктиром, питается от того же источника, что и основной стабилизатор. Обозначив $\Delta E_0 = h \Delta U_1$ и включив эту э. д. с. последовательно с сопротивлением R_0 на рис. 23-9,б, можно показать, что коэффициент стабилизации будет в $1 + \frac{hr_d R_i}{R_0 R_{\infty}}$ раз меньше, чем (23-46).

Многокаскадные стабилизаторы. Главным недостатком однокаскадной схемы, рассмотренной в предыдущем разделе, является

сравнительно высокое выходное сопротивление. Это обстоятельство, как уже отмечалось, ограничивает выходной ток.

В двухкаскадном стабилизаторе, показанном на рис. 23-11, транзистор T_1 является регулирующим элементом, а транзистор T_2 — сравнивающим и усилительным. Следовательно, в данном случае $I_p = I_{к1}$; $R_y = r_d + R_{вх2}$ и $K_i = \beta_2 \beta_1$, где $R_{вх2}$ — входное сопротивление усилительного каскада ОЭ:

$$R_{вх2} = r_{с2} + (1 + \beta_2) r_{э2}.$$

Используя формулу (23-17), получаем характеристическое сопротивление, близкое к выходному сопротивлению стабилизатора. Запишем их в следующей форме:

$$R_{вых} \approx R_{\infty} = \frac{r_d + r_{с2} + (1 + \beta_2) r_{э2}}{1 + \beta_1 \beta_2}, \quad (23-48)$$

Рис. 23-11. Двухкаскадный последовательный стабилизатор со стабитроном в базовой цепи.

где $r_{э2} = \varphi_T / I_{э2}$. Например, если $I_{э2} = 10$ мА, $r_{с2} = 50$ Ом; $\beta_1 = \beta_2 = 30$ и $r_d = 10$ Ом, то $R_{вых} \approx 0,15$ Ом.

Как видим, выигрыш по сравнению с однокаскадной схемой весьма значительный. Соответственно возрастает коэффициент стабилизации. Согласно (23-43в) при $R_i = r_{к1}^*$ и $U_{p.мин} = 0$ получаем:

$$K \approx \frac{(1 - \delta_{н1}) r_{к1}^*}{R_{\infty}} \approx \frac{\beta_2 (1 - \delta_{н1}) r_{к1}}{r_d + r_{с2} + (1 + \beta_2) r_{э2}}. \quad (23-49)$$

Если параметры имеют указанные выше значения и $r_{к1} = 5$ кОм, то $K \approx 1000$.

Токоотводящий резистор R_0 выполняет примерно ту же функцию, что и на рис. 23-9, а его сопротивление рассчитывается по формуле

$$R_0 = \frac{E_0 - U_{б1}}{I_{б1 макс} + I_{к2 мин}}, \quad (23-50)$$

где $U_{б1} \approx U_2$ и $I_{б1 макс} \approx I_{э1 макс} / (1 + \beta_1)$.

Напряжение E_0 должно быть хорошо стабилизировано. В противном случае, как и в однокаскадной схеме, уменьшается коэффициент стабилизации. Это уменьшение можно оценить с помощью (23-44), так как случай $E_0 = f(U_1)$ равен силён наличию цепи прямой связи (рис. 23-8). Подставив в (23-44) значения $R_i = r_{к1}^*$; $R_{y4} = \frac{1}{h} R_0$ и $K_{i4} = 1 + \beta_1$ (где $h = \Delta E_0 / \Delta U_1$ — коэффициент, характеризующий зависимость E_0 от входного напряжения¹⁾), получим, что коэффициент стабилизации уменьшается в $1 + h r_{к1} / R_0$ раз. Следовательно, стабильность напряжения E_0 нужно обеспечивать из условия $h < R_0 / r_{к1}$. Если $h = 1$, т. е.

¹ Выражение $R_{y4} = \frac{1}{h} R_0$ получается, если приравнять ток $\Delta I_{y4} = \Delta U_1 / R_{y4}$ (рис. 23-8) току $\Delta I_{R_0} = \Delta E_0 / R_0$ (рис. 23-11) при $U_{б1} \approx U_2 = \text{const}$.

если резистор R_0 питать непосредственно от входного напряжения, то коэффициент стабилизации уменьшается в $1 + r_{к1}/R_0$ раз и выражается формулой (23-49), в которой $r_{к1}$ следует заменить гораздо меньшей величиной R_0 . При малых токах значения K могут составлять 10—20 и менее, а значит, такая схема оказывается практически непригодной.

Обычно минимальный рабочий ток стабилитрона больше базового тока в транзисторе T_2 . Поэтому приходится вводить дополнительное смещение в виде токоподводящей цепи E_d, R_d , показанной на рис. 23-11 пунктиром. Параметры этой цепи выбираются из очевидного соотношения

$$I_d - I_{б1} = I_{R_d} \approx E_d / R_d. \quad (23-51)$$

Чтобы избежать применения токоподводящей цепи, включают стабилитрон не в базовую, а в эмиттерную цепь усилительного транзистора (рис. 23-12); эмиттерный ток обычно достаточен для питания стабилитрона: $I_{э2\text{мин}} > I_{д.\text{мин}}$ (в противном случае следует питать стабилитрон от выходного напряжения через балластный резистор, см. рис. 23-13). Недостатком эмиттерного включения является повышенное значение входного сопротивления усилителя

$R_y = R_{вх2} = r_{б2} + (1 + \beta_2)(r_{э2} + r_d)$ и соответственно характеристического и выходного сопротивлений

$$R_{вых} \approx R_\infty = \frac{(r_{э2} + r_d)(1 + \beta_2) + r_{б2}}{1 + \beta_1\beta_2}. \quad (23-52)$$

Увеличение R_∞ приводит к уменьшению коэффициента стабилизации:

$$K = \frac{(1 + \beta_1\beta_2)(1 - \delta_{н1})r_{к1}^*}{r_{б2} + (r_d + r_{э2})(1 + \beta_2)} \approx \frac{(1 + \delta_{н1})r_{к1}}{r_d + r_{э2} + r_{б2}/\beta_2}. \quad (23-53)$$

Полагая значения параметров прежними, получаем $R_\infty \approx 0,45 \text{ Ом}$ и $K \approx 300$, т. е. в 3 раза хуже, чем при базовом включении. Большее значение R_∞ при прочих равных условиях приводит к меньшему допустимому току нагрузки.

Двухкаскадные стабилизаторы с параметрами $R_{вых} = 0,1 \div 0,5 \text{ Ом}$; $K = 200 \div 500$ и рабочими токами $0,2 \div 0,5 \text{ А}$ являются типовыми для многих устройств. В случае больших токов (мощностей) и повышенных требований к стабильности необходимо дальнейшее уменьшение характеристического сопротивления путем увеличения коэффициента усиления K_i . Это достигается либо применением многокаскадных усилителей вместо одного каскада T_2 [173], либо применением составного транзистора в качестве регу-

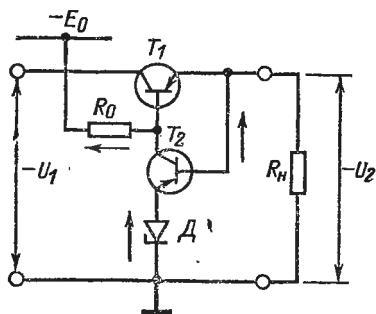


Рис. 23-12. Двухкаскадный последовательный стабилизатор со стабилитроном в эмиттерной цепи.

лирующего транзистора T_1 . Последний путь иллюстрируется схемой на рис. 23-13.

Здесь составным транзистором (см. § 4-11) является комбинация транзисторов T_1 , T_2 и T_3 . Как известно, составные транзисторы характеризуются огромными значениями β_Σ ; в данном случае $\beta_\Sigma \approx \beta_1\beta_2\beta_3$ и может составлять десятки тысяч вместо 30—50 для одиночного транзистора. Заменяя в формулах (23-48), (23-49) и (23-52), (23-53) индексы 2 индексами 4 и подставляя β_Σ вместо β , легко

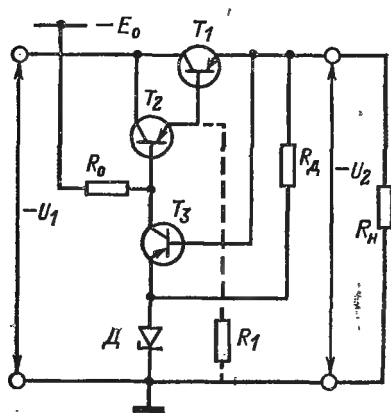


Рис. 23-13. Последовательный стабилизатор с составным регулирующим транзистором.

удет заметить, что приближение $U_6 \approx U_2$ в формулах (23-50) и (23-54) может привести к заметным ошибкам, так как напряжения U_{63} при больших токах превышают 1 В у германиевых транзисторов и 2 В у кремниевых.

убедиться, что выходные сопротивления уменьшаются, а коэффициенты стабилизации увеличиваются в $\beta_\Sigma\beta_3$ раз. В такой схеме значение $R_{\text{вых}}$ может составлять сотые и тысячные доли ома. Количество транзисторов в составном транзисторе диктуется конкретными требованиями к стабилизатору.

Резисторы $R_{01,2}$, показанные пунктиром, призваны обеспечить нормальные рабочие режимы составляющих транзисторов; сопротивления этих резисторов выбираются из соотношения

$$R_{6v} = \frac{U_{6v}}{I_{\Sigma(v+1)} - I_{6v}}, \quad (23-54)$$

где токами I_{Σ} и I_6 задаются с учетом тока нагрузки и коэффициентов β . Сопротивление R_0 оценивается из формулы (23-50) при замене индексов 1 и 2 на 3 и 4. След-

23-4. СРАВНЕНИЕ ПАРАЛЛЕЛЬНЫХ И ПОСЛЕДОВАТЕЛЬНЫХ СТАБИЛИЗАТОРОВ

Из соотношения (23-6а) следует, что параллельные стабилизаторы не чувствительны к перегрузкам по току, так как с увеличением тока I_n ток регулирующего элемента уменьшается. При токах I_n , заметно больших расчетного значения $I_{n, \text{макс}}$, регулирующий транзистор запирается. При коротком замыкании на выходе напряжение U_1 полностью падает на балластном сопротивлении R_0 и регулирующий транзистор оказывается вне опасности. Последовательные стабилизаторы, как следует из соотношения (23-32а), чувствительны к перегрузкам, поскольку ток нагрузки и ток регулирующего элемента возрастают одновременно и в равной степени. При токах $I_n > I_{n, \text{макс}}$ усилительный и опорный элементы оказываются запертыми, а регулирующий транзистор работает с максимальным базовым током, определяемым величиной токоотводящего сопротивления и разностью потенциалов $U_1 - U_2$. Короткое замыкание на выходе ($U_2 = 0$) увеличивает базовый ток регулирующего транзистора и напряжение на нем в отношении $U_1/(U_1 - U_2)$, т. е. обычно в несколько раз. При этом рассеиваемая

мощность возрастает на порядок и больше и транзистор неизбежно выходит из строя. Этот недостаток последовательных стабилизаторов заставляет дополнять их схему тем или иным типом защиты [176].

Такая защита при заданном превышении нагрузочного тока над расчетным значением $I_{н. макс}$ либо быстро снимает напряжение питания, либо резко уменьшает ток регулирующего транзистора, отключая его базу от токоотводящего резистора. Оригинальный и перспективный метод триггерной защиты предложен в работе [176].

Сравнивая выражения (23-9) и (23-37), приходим к выводу, что при одном и том же выходном напряжении и обычных значениях допусков $\delta_1 = 0,1 \div 0,2$ в последовательных стабилизаторах требуется менее высоковольтный транзистор, чем в параллельных. Однако этот вывод не учитывает аварийной ситуации, когда на регулирующем транзисторе может в течение короткого времени действовать полное напряжение питания. Поэтому практически в обоих типах стабилизаторов ориентируются на одно и то же условие $U_{р. доп} \geq U_{1 макс}$. Усилительные транзисторы выбирают из того же условия, что и регулирующий элемент.

Сравнивая правые части неравенств (23-16) и (23-40), легко установить, что при одном и том же токе нагрузки в параллельных стабилизаторах требуются примерно вдвое более сильноточные транзисторы, чем в последовательных. По мощности разница получается еще больше. Однако при наличии «гасящего» сопротивления в параллельных стабилизаторах (см. с. 626) разница в допустимой мощности регулирующих элементов делается практически несущественной.

Коэффициент полезного действия у последовательных стабилизаторов согласно (23-39) зависит от напряжения $U_{р. мин}$, которое не входит в выражение (23-13) для параллельных стабилизаторов. Поэтому однозначное сравнение, строго говоря, невозможно. Однако если положить $U_{р. мин} = 0$ и учесть, что в знаменателе (23-13) неявно стоит слагаемое U_2/R_i («неизменная нагрузка», см. сноску на с. 623), то придем к выводу, что *к. п. д. у последовательных стабилизаторов несомненно выше, чем у параллельных*. Учет напряжения $U_{р. мин}$, если оно не превышает 2—3 В, не меняет этого важного вывода.

Для дифференциальных параметров определяющим является характеристическое сопротивление R_{∞} . Согласно (23-17) оно в

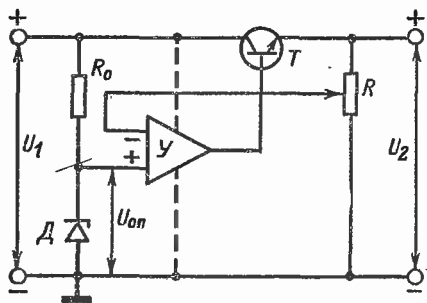


Рис. 23-14. Последовательный стабилизатор с использованием интегральной микросхемы усилителя (см. узел 2 на рис. 23-2).

принципе может быть одинаковым у обоих типов стабилизаторов, если обеспечить одинаковые значения R_y и K_i .

Необходимо подчеркнуть, что в современных стабилизаторах усилительная часть, как правило, реализуется в виде одной детали — интегрального усилителя постоянного тока (рис. 23-14). Следовательно, характеристическое сопротивление определяется именно параметрами этой детали, а не типом стабилизатора.

Коэффициенты стабилизации помимо R_∞ зависят ещё от R_0 и R_i [см. (23-19a) и (23-43a)]. Практически всегда $R_i \gg R_{0\text{опт}}$, что говорит в пользу последовательных стабилизаторов.

Однако при решении конкретных задач параллельные стабилизаторы могут быть практически равноценным, а с учетом перегрузочной способности — даже оптимальным вариантом.

23-5. ОСОБЕННОСТИ ПРАКТИЧЕСКИХ СХЕМ

Регулировка выходного напряжения. Рассмотренные выше стабилизаторы обеспечивают выходное напряжение, близкое к напряжению стабилитрона и имеющее отрицательную полярность. На практике часто требуется иметь напряжение, отличное от U_d ,

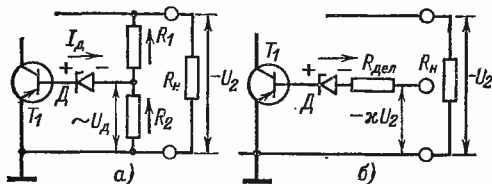


Рис. 23-15. Регулировка выходного напряжения с помощью делителя.

а — принципиальная схема; б — преобразованная схема.

регулируемое (ступенями или плавно), а также имеющее положительную полярность. При всех этих вариантах основная часть схемы мало меняется, и мы остановимся только на тех изменениях, которые необходимы в том или ином случае.

Путем последовательного включения стабилитронов (согласного и встречного) можно обеспечить практически любое напряжение, не превышающее допустимого значения для регулирующих и усилительных транзисторов. Однако количество используемых стабилитронов не должно быть большим, так как сопротивление составного стабилитрона больше, чем у каждого из его компонентов, а это ухудшает параметры схемы. Кроме того, встречное включение, как легко убедиться, требует применения токоотводящих резисторов в точках соединения стабилитронов. Общим недостатком такого решения является невозможность плавной регулировки выходного напряжения.

На рис. 23-15,а показан наиболее распространенный способ повышения выходного напряжения с помощью делителя R_1 , R_2 . Этот способ в равной степени относится к последовательной и параллельной схемам, а также к эмиттерному включению стабилитрона (см. рис. 23-12); в последнем случае нужно заменить ток I_d током

$I_6 \approx I_d/\beta$. Пренебрегая напряжением U_{96} , получаем следующее соотношение:

$$U_2 = U_d + \left(I_d + \frac{U_d}{R_2} \right) R_1. \quad (23-55)$$

Отсюда видно, что сопротивление R_2 , вообще говоря, не является необходимым. При $R_2 = \infty$ можно обеспечить желательную разность потенциалов $U_2 - U_d$ за счет протекания тока I_d по сопротивлению R_1 . Однако поскольку ток I_d невелик, для получения разности $U_2 - U_d$ от нескольких вольт и выше пришлось бы ставить сопротивления R_1 порядка сотен и тысяч ом. Это привело бы к значительному увеличению сопротивления R_y и резко ухудшило бы параметры стабилизатора. Сопротивление R_2 , которое обеспечивает дополнительные ток через R_1 , позволяет уменьшить величину последнего. Это уменьшение будет существенным при условии

$$I_{R_2} \approx \frac{U_d}{R_2} \gg I_d$$

т. е. если

$$R_2 \ll \frac{U_d}{I_d}. \quad (23-56)$$

Тогда токи в обоих сопротивлениях почти одинаковы и значительно превышают ток базы.

Считая условие (23-56) выполненным, получим вместо (23-55) следующее соотношение:

$$U_2 \approx U_d \left(1 + \frac{R_1}{R_2} \right). \quad (23-57)$$

Сделав сопротивления делителя переменными, можно плавно менять выходное напряжение от значения, близкого к U_d , до весьма больших значений.

Влияние делителя на параметры стабилизатора заслуживает специального рассмотрения. Преобразуем делитель по теореме об эквивалентном генераторе. Тогда получается схема, показанная на рис. 23-15,б, в которой $R_{\text{дел}} = R_1 \parallel R_2$; $R_2/(R_1 + R_2) = \kappa$. Приращение входного тока усилителя, которое в отсутствие делителя выражалось бы формулой

$$\Delta I_y = \Delta U_2 / (r_d + R_y),$$

теперь имеет вид:

$$\Delta I_y = \kappa \Delta U_2 / (r_d + R_y + R_{\text{дел}}).$$

Ясно, что наличие делителя приводит к уменьшению приращений входного тока при одном и том же ΔU_2 . Это равносильно увеличению сопротивления $R_y = \Delta U_2 / \Delta I_y$ и соответственно характеристического сопротивления R_{∞} . Поэтому *параметры стабилизатора с делителем всегда хуже, чем без делителя*. Учитывая приведенные выражения для токов ΔI_y , можно записать:

$$\frac{K}{K_0} = \frac{R_{\text{вых}0}}{R_{\text{вых}}} = \frac{\kappa}{1 + R_{\text{дел}}/(r_d + R_y)}, \quad (23-58)$$

где индекс 0 означает отсутствие делителя.

Выражение (23-58) подтверждает, что сопротивления делителя желательно делать возможно меньшими, стремясь в пределе к неравенству

$$R_{\text{дел}} \ll r_d + R_y. \quad (23-59)$$

Практически условие (23-59) редко выполнимо; оно приводит к таким токам в делителе, которые сравнимы с током нагрузки, что резко снижает к. п. д. и требует излишне мощных регулирующих транзисторов. По этой причине использование делителей в маломощных однокаскадных стабилизаторах не всегда возможно. Заметим еще, что наличие делителя в последовательных стабилизаторах автоматически решает проблему шунта, обеспечивающего минимальный ток регулирующего элемента при холостом ходе [см. замечания к (23-34)].

Из формулы (23-58) видно, что даже в предельном случае, при $R_{\text{дел}} = 0$, параметры схемы с делителем в $1/\kappa$ раз хуже, чем параметры без делителя, т. е. во столько раз, во сколько напряжение U_2 превышает напряжение U_d . Следовательно, при регулировке выходного напряжения параметры стабилизатора неизбежно меняются.

Они имеют оптимальные значения при минимальном выходном напряжении и ухудшаются с ростом напряжения. Поэтому при расчете параметров стабилизатора нужно ориентироваться на худший случай — максимальное выходное напряжение.

При эмиттерном включении стабилитрона неравенство, аналогичное (23-59), имеет вид:

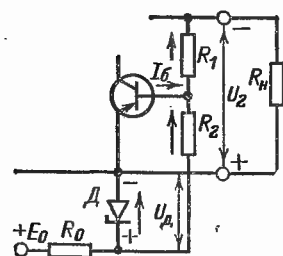
$$R_{\text{дел}} \ll (1 + \beta) r_d + R_{\text{вх}}. \quad (23-60)$$

Это неравенство легче выполнить на практике, особенно в мощных стабилизаторах, вследствие чего ухудшение параметров по сравнению с такой же схемой без делителя получается сравнительно небольшим.

Рис. 23-16. Использование делителя со стабилитронным смещением для получения малых выходных напряжений.

Для получения малых напряжений, регулируемых как по величине, так и по полярности, обычно используют встречное включение двух стабилизаторов с разными выходными напряжениями. Результирующее выходное напряжение $U_2 = U_{21} - U_{22}$ может быть сколь угодно малым и может регулироваться в пределах $\pm U_{\text{макс}}$. Выходное сопротивление такой системы, разумеется, равно сумме $R_{\text{вых1}} + R_{\text{вых2}}$, а коэффициент стабилизации получается хуже, чем у каждого из стабилизаторов, и уменьшается с уменьшением результирующего выходного напряжения.

На рис. 23-16 показан еще один способ, позволяющий получить малые выходные напряжения. Если выполняется условие (23-56), то через сопротивление R_1 проходит заданный ток $I_{R2} \approx U_d / R_2$. Поэтому, меняя значение R_1 , можно получить как весьма большие, так и весьма малые напряжения U_2 , близкие к нулю. Практически $U_{2\text{мин}} \approx U_{96} \approx 0,5$ В. Изменение полярности выходного напряжения в такой схеме, конечно, невозможно.



Для получения положительной полярности выходного напряжения в параллельных стабилизаторах, в частности диодных, достаточно изменить точку заземления и переместить балластное сопротивление в провод, соединенный с положительным входным зажимом. В последовательных стабилизаторах возможны еще два пути: либо использовать транзисторы $n-p-n$ во всех каскадах при сохранении общей структуры схемы, либо подключить нагрузку в коллекторную цепь регулирующего транзистора $p-n-p$ (рис. 23-17). Схема, показанная на рис. 23-17, имеет практически те же параметры, что и схема на рис. 23-11, так как эмиттерный и коллекторный токи транзистора почти одинаковы и почти одинаково зависят от управляющего тока базы. Однако изменившаяся полярность выходного напряжения заставляет использовать в усиленном каскаде транзистор типа $n-p-n$ и несколько изменить способ включения стабилитрона [177]. Интересно отметить, что сочетание транзисторов $p-n-p$ и $n-p-n$, вообще говоря, исключает необходимость в токоотводящем сопротивлении (R_0 на рис. 23-11).

Включение амперметра, измеряющего ток нагрузки, между выходом регулирующего транзистора и цепью обратной связи характерно для любого стабилизатора. В противном случае (если включить прибор после цепи обратной связи) сопротивление амперметра непосредственно складывается с выходным сопротивлением схемы и существенно увеличивает его.

Температурная стабильность. Во всех рассмотренных стабилизаторах, как параллельных, так и последовательных, выходное напряжение либо равно сумме

$$U_2 = U_d + U_{эб}, \quad (23-61)$$

либо пропорционально этой сумме. Поэтому временной и температурный дрейфы выходного напряжения определяются изменениями компонентов U_d и $U_{эб}$ при неизменных значениях U_1 и R_n . Временной дрейф напряжения U_d , как отмечалось в § 3-3, практически отсутствует, а временной дрейф напряжения $U_{эб}$ носит хаотический характер и во многом зависит от качества транзистора. Соответствующие данные приводились в § 17-3.

Зависимости напряжений U_d и $U_{эб}$ от температуры при заданных токах характеризуют температурными чувствительностями напряжения, о которых говорилось в § 2-8 и 3-3. Напомним, что температурная чувствительность стабилитронов обычно положительна, возрастает с увеличением номинала U_d и увеличивается с ростом тока I_d . Температурная чувствительность эмиттерного перехода отрицательна при малых токах и уменьшается с ростом тока $I_э$. Поэтому результирующая температурная чувствительность выходного напряжения зависит как от режима транзистора, так и от номинала U_d .

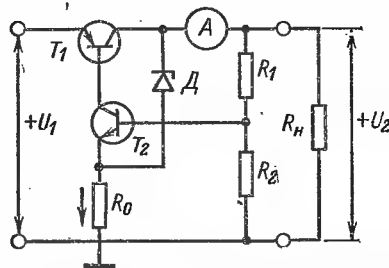


Рис. 23-17. Последовательный стабилизатор с положительной полярностью выходного напряжения.

Практически положительная составляющая температурной чувствительности всегда превалирует, т. е. выходное напряжение растет с температурой. Суммарная температурная чувствительность обычно лежит в пределах 2—5 мВ/°С. Если это значение неприемлемо, следует использовать составной опорный элемент, в котором сочетаются прямые и обратные включения стабилитронов. С помощью токоотводящих сопротивлений можно регулировать ток в каждом из стабилитронов последовательной цепочки, добиваясь оптимального результата. В транзисторных стабилизаторах можно часть стабилитронов включать в эмиттерную цепь, а часть — в базовую, используя различие токов в этих цепях для получения необходимых значений температурной чувствительности. Составные опорные элементы позволяют в ряде случаев снизить суммарную температурную чувствительность до 0,1 мВ/°С. Однако взаимная компенсация температурной чувствительности имеет место лишь в сравнительно узком диапазоне тока I_d , что нужно учитывать при конструировании схемы. Общее сопротивление r_d комбинированного опорного элемента равно сумме сопротивлений отдельных стабилитронов.

Изменения температуры не только обуславливают дрейф выходного напряжения, но и влияют на параметры стабилизатора K и $R_{\text{вых}}$. Так, формулы (23-26) и (23-48) показывают, что характеристическое сопротивление зависит, в частности, от величины β , которая меняется с температурой. Учитывая данные, приведенные в § 4-6 (см. рис. 4-20), можно произвести приближенную оценку зависимости $R_{\text{вых}}(T)$. Обычно в диапазоне $\pm 60^\circ \text{C}$ выходное сопротивление изменяется на 50% и больше, уменьшаясь с ростом температуры.

Коэффициент стабилизации в параллельных схемах меняется обратно пропорционально характеристическому сопротивлению и, следовательно, увеличивается с ростом температуры. В последовательных схемах коэффициент стабилизации согласно (23-49) зависит еще и от величины r_k , которая меняется с температурой по сложному закону. Соответственно более сложной будет зависимость $K(T)$.

При отрицательных температурах значение K заметно уменьшается в обоих типах стабилизаторов.

Зависимость параметров от частоты и режима. Выходные сопротивления транзисторных стабилизаторов, особенно многокаскадных, чрезвычайно малы. Однако этот вывод относится лишь к статическим величинам $R_{\text{вых}}$. При скачкообразных изменениях тока нагрузки, а также при наличии в нем высокочастотных составляющих приходится считаться с комплексным характером сопротивления $Z_{\text{вых}}$ [176]. Например, при скачкообразном изменении тока I_n коэффициент β у транзисторов в первый момент равен нулю. Соответственно в формуле (23-17) будет $K_i(0) = 0$, и тогда начальное выходное сопротивление

$$R_{\text{вых}}(0) \approx R_{\infty}(0) \approx r_d + r_b + r_a.$$

значении $R_{\text{вых}}(0)$ может в десятки раз превышать установившееся значение $R_{\text{вых}}$. Поэтому выходное напряжение в первый момент после скачка может существенно измениться и восстановится лишь спустя некоторое время, определяемое постоянными времени τ_β . Во избежание таких переходных процессов выход стабилизатора шунтируется конденсатором достаточно большой емкости. Эту емкость можно выбирать из условия

$$CR_{\text{вых.ср}} \gg \tau_0, \quad (23-62)$$

где τ_0 — эквивалентная постоянная времени (которую в первом приближении можно считать суммой постоянных времени τ_β каждого из транзисторов), а $R_{\text{вых.ср}}$ — усредненное (по интервалу переходного процесса) выходное сопротивление. Весьма грубым приближением можно считать $R_{\text{вых.ср}} = \sqrt{R_{\text{вых}}(0) R_{\text{вых}}}$. При усредненном выходном сопротивлении, примерно равном 0,1 Ом, и величине τ_0 около 10 мкс необходимое значение емкости получается весьма большим — сотни микрофард.

Коэффициент стабилизации как функция R_∞ тоже является комплексной величиной, модуль которой уменьшается с ростом частоты пульсаций и скорости изменения входного напряжения. Однако скачкообразные изменения напряжения U_1 маловероятны, так как стабилизатор обычно питается от выпрямителя с хорошим фильтром. Конденсатор C , включенный на выходе стабилизатора, уменьшает влияние и этих изменений, если они имеют место.

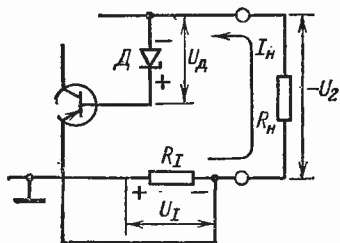


Рис. 23-18. Происхождение положительной обратной связи по току.

Наличие нескольких фазосдвигающих звеньев в цепи обратной связи при достаточно большом коэффициенте усиления может обусловить колебательный характер переходного процесса и даже привести к самовозбуждению схемы. Увеличение емкости C не всегда достаточно для ликвидации этих нежелательных явлений. Поэтому иногда приходится дополнять схему специальными стабилизирующими цепочками, руководствуясь общими принципами, известными из теории усилителей (см., например, [62], § 5-5).

Еще одной причиной самовозбуждения может быть паразитная обратная связь по току нагрузки. На рис. 23-18 показана часть стабилизатора с сопротивлением R_I , которое обуславливает такую обратную связь. Вообще говоря, сопротивление R_I позволяет уменьшить выходное сопротивление схемы, но может сделать его и отрицательным. В самом деле, если $R_I = 0$, то выходное напряжение U_2 , как известно, уменьшается с ростом тока I_H . Если же $R_I \neq 0$, то напряжение $U_I = I_H R_I$, складываясь с опорным напряжением U_d , стремится увеличить напряжение U_2 . При условии $R_I = R_{\text{вых}}$ теоретически должна иметь место компенсация обоих явлений, что равносильно нулевому результирующему сопротивлению. При условии $R_I > R_{\text{вых}}$ результирующее сопротивление будет отрицательным, а это может привести к самовозбуждению схемы. Учитывая, что в ряде случаев

роль сопротивления R_1 может играть сопротивление монтажных проводов, желательно при монтаже стабилизаторов с очень малым выходным сопротивлением производить пайку всех концов, идущих к выходным зажимам, в двух узловых точках.

Важнейший параметр всякого стабилизатора — характеристическое сопротивление R_{∞} — зависит от режима работы транзисторов и, следовательно, от значений входного напряжения и нагрузочного тока. Последняя зависимость играет главную роль, если ток нагрузки меняется в широких пределах. Как известно, коэффициент передачи β и сопротивление эмиттерного перехода r_e уменьшаются с увеличением рабочего тока транзистора, причем зависимость $r_e(I_e)$ более сильная. Поэтому характер функции $R_{\infty}(I_n)$, определяющей параметры K и $R_{\text{вых}}$, во многом зависит от относительной роли члена $(1 + \beta)r_e$ в числителях выражений (23-26) и (23-48). Роль этого члена особенно существенна при малых токах первого усиленного транзистора, т. е. при полной нагрузке стабилизатора.

Функции $R_{\text{вых}}(I_n)$ и $K(I_n)$ в общем случае могут быть немонотонными. Однако в последовательных стабилизаторах, в которых важное влияние на коэффициент стабилизации оказывают не столько параметры r_e и β , сколько сопротивление r_k регулирующего транзистора, значение K , как правило, уменьшается с увеличением нагрузочного тока.

Изменения входного напряжения влияют на параметры стабилизаторов постольку, поскольку в последовательных схемах при этом меняется напряжение на регулирующем транзисторе, а в параллельных — его ток.

23-6. МЕТОДИКА РАСЧЕТА СТАБИЛИЗАТОРОВ

Энергетический расчет стабилизаторов, т. е. выбор регулирующего элемента (по напряжению, току и мощности), определение входного напряжения и оптимального балластного сопротивления осуществляются по общим формулам, приведенным в § 23-2 и 23-3, с учетом к. п. д. и минимальных величин $I_{p.\text{мин}}$, $U_{p.\text{мин}}$, $I_{n.\text{мин}}$, которыми задаются из тех или иных соображений.

При выборе стабилитронов или их комбинаций (составных опорных элементов) руководствуются, конечно, значением выходного напряжения ($U_d \leq U_{2\text{мин}}$), а также желательным значением температурного дрейфа (см. § 23-5).

Что касается усилителя, то его главная функция — обеспечить необходимое значение характеристического сопротивления и тем самым дифференциальных параметров $R_{\text{вых}}$ и K . Последние характеризуют качество стабилизации, и соответствующие расчеты полезно рассмотреть несколько подробнее.

На практике чаще всего встречаются два варианта расчетов: 1) по допустимым отклонениям выходного напряжения в зависимости от изменений тока нагрузки и входного напряжения: $\delta_{2\text{доп}}^I$ и $\delta_{2\text{доп}}^U$, [см. (23-4)]; 2) по желательным

величинам выходного сопротивления стабилизации K . Строго говоря, первый вариант является основным, поскольку параметры $R_{\text{вых}}$ и K в конце концов призваны обеспечить допустимые значения δ_2^I и δ_2^U .

Рассмотрим эти два варианта.

1. Заданы значения $\delta_{2\text{доп}}^I$ и $\delta_{2\text{доп}}^U$. Подставим в выражение (23-46) $R_{\text{вых}} = R_{\infty}$ и положим $\delta_2^I \leq \delta_{2\text{доп}}^I$; тогда получим:

$$R_{\infty} \leq \frac{\delta_{2\text{доп}}^I U_2}{I_{\text{н. макс}}} = \delta_{2\text{доп}}^I R_{\text{н. мин}}. \quad (23-63)$$

Подставим в выражение (23-4в) значение K из (23-43в) или (23-19в) и положим $\delta_2^U \leq \delta_{2\text{доп}}^U$, тогда получим:

для последовательной схемы

$$R_{\infty} \leq \frac{\delta_{2\text{доп}}^U (1 - \delta_{\text{н1}}) U_2 R_i}{\delta_1 (U_2 + U_{\text{р. мин}})}; \quad (23-64а)$$

для параллельной схемы

$$R_{\infty} \leq \frac{\delta_{2\text{доп}}^U (1 - \delta_{\text{н1}}) U_2 R_0}{\delta_1 (U_2 + I_{1\text{ мин}} R_0)}. \quad (23-64б)$$

Здесь для последовательного стабилизатора сопротивление R_i может считаться известным.¹ Для параллельного стабилизатора сопротивление R_0 , вообще говоря, величина произвольная; при оптимальном режиме, подставляя (23-12), получаем:

$$R_{\infty} \leq \frac{\delta_{2\text{доп}}^U U_2}{I_{1\text{ мин}}} \frac{\sqrt{\delta_{\text{н1}}} (1 - \sqrt{\delta_{\text{н1}}})}{\delta_1}. \quad (23-64в)$$

Очевидно, что из двух значений R_{∞} , определяемых неравенствами (23-63) и (23-64), следует выбрать меньшее. Поскольку R_{∞} согласно (23-17) зависит только от параметров R_y и K_i и поскольку параметр R_y трудно менять в широких пределах, необходимое значение R_{∞} обеспечивают в первую очередь за счет параметра K_i , рассчитывая усилительный элемент стабилизатора на усиление K_i/β_p , где β_p — коэффициент усиления регулирующего элемента.

2. Заданы параметры $R_{\text{вых}}$ и K . Тогда из определения коэффициента K , полагая $R_{\infty} = R_{\text{вых}}$ и подставляя значения U_i , нетрудно получить:

для последовательного стабилизатора

$$R_i \geq \frac{K (U_2 + U_{\text{р. мин}})}{(1 - \delta_{\text{н1}}) U_2} R_{\text{вых}}; \quad (23-65а)$$

для параллельного стабилизатора

$$R_0 \geq \frac{K U_2}{(1 - \delta_{\text{н1}}) U_2 - K I_{1\text{ мин}} R_{\text{вых}}} R_{\text{вых}}. \quad (23-65б)$$

В последнем выражении, конечно, подразумевается, что $K < K_{\infty}$, где K_{∞} определяется формулой (23-20). В противном случае правая часть (23-65б) получается отрицательной.

Удовлетворить неравенствам (23-65) далеко не всегда удастся, так как сопротивление R_i задано вместе с регулирующим транзистором, а сопротивление R_0 , выбранное из (23-65б), может оказаться намного больше оптимального. Поэтому чаще всего приходится «подгонять» $R_{\text{вых}}$ (в сторону уменьшения) под желательные значения R_i и R_0 . Коэффициент усиления K_i рассчитывается, как и в первом варианте, по значению $R_{\text{вых}} = R_{\infty}$ согласно (23-17).

¹ Поскольку коллекторное сопротивление транзистора зависит от тока, следует использовать значение R_i , соответствующее максимальному току нагрузки.

ТАБЛИЦА СООТВЕТСТВИЯ ОБОЗНАЧЕНИЙ ЭЛЕКТРИЧЕСКИХ ПАРАМЕТРОВ
БИПОЛЯРНЫХ ТРАНЗИСТОРОВ, ПРИНЯТЫХ В КНИГЕ И ПО ГОСТ 20003-74

Электрический параметр	Буквенное обозначение	
	в книге	по ГОСТ 20003-74
Обратный ток коллектора	$I_{к0}$	$I_{КБО}$
Обратный ток эмиттера	$I_{э0}$	$I_{ЭБО}$
Напряжение насыщения коллектор — эмиттер		$U_{КЭнас}$
Напряжение насыщения база — эмиттер		$U_{БЭнас}$
Плавающее напряжение эмиттер — база		$U_{ЭБпл}$
Пробивное напряжение эмиттер — база		$U_{ЭБОпроб}$
Пробивное напряжение коллектор — база		$U_{КБОпроб}$
Входное сопротивление биполярного транзистора в режиме малого сигнала	h_{11}	h_{11}^*
Коэффициент обратной связи по напряжению биполярного транзистора в режиме малого сигнала	$h_{12}, \mu_{эк}$	h_{12}^*
Коэффициент передачи тока биполярного транзистора в режиме малого сигнала	α, β	h_{21}^*
Модуль коэффициента передачи тока биполярного транзистора на высокой частоте		$ h_{21э} $
Выходная полная проводимость биполярного транзистора в режиме малого сигнала при холостом ходе		h_{22}^*
Входное сопротивление биполярного транзистора в схеме с общим эмиттером в режиме большого сигнала		$h_{11э}$
Статический коэффициент передачи тока биполярного транзистора в схеме с общим эмиттером	$\bar{\beta}$	$h_{21э}$
Входная полная проводимость биполярного транзистора в режиме малого сигнала		Y_{11}^*
Полная проводимость обратной передачи биполярного транзистора в режиме малого сигнала		Y_{12}^*
Полная проводимость прямой передачи биполярного транзистора в режиме малого сигнала		Y_{21}^*
Выходная полная проводимость биполярного транзистора в режиме малого сигнала при коротком замыкании		Y_{22}^*
Статическая крутизна прямой передачи в схеме с общим эмиттером		$Y_{21э}$

Электрический параметр	Буквенное обозначение	
	в книге	по ГОСТ 20003-74
Входная емкость биполярного транзистора		C_{11}^*
Выходная емкость биполярного транзистора		C_{22}^*
Предельная частота коэффициента передачи тока биполярного транзистора	f_T	f_{h21}
Граничная частота коэффициента передачи тока в схеме с общим эмиттером	f_β	$f_{гр}$
Максимальная частота генерации биполярного транзистора	$f_{ген. макс}$	$f_{мах}$
Коэффициент шума биполярного транзистора	$F_{ш}$	$K_{ш}$
Коэффициент насыщения биполярного транзистора	N	$K_{нас}$
Коэффициент усиления по мощности биполярного транзистора	K_p	$K_{ур}$
Коэффициент полезного действия коллектора		η_k
Время задержки для биполярного транзистора	t_3	$t_{зд}$
Время нарастания для биполярного транзистора	t_Φ^+	$t_{нр}$
Время рассасывания для биполярного транзистора	t_p	$t_{рас}$
Время спада для биполярного транзистора	t_Φ^-	$t_{сп}$
Время включения биполярного транзистора		$t_{вкл}$
Время выключения биполярного транзистора		$t_{выкл}$
Сопротивление базы биполярного транзистора	r_b	r'_b
Постоянный ток коллектора	I_k	I_k
Постоянный ток эмиттера	I_ε	I_ε
Постоянный ток базы	I_b	I_B
Постоянный ток коллектора в режиме насыщения	$I_{к. н}$	$I_{Кнас}$
Постоянный ток базы в режиме насыщения	$I_{б. н}$	$I_{Бнас}$
Постоянное напряжение эмиттер — база	$U_{эб}$	$U_{ЭБ}$
Постоянное напряжение коллектор — база	$U_{кб}$	$U_{КБ}$
Постоянное напряжение коллектор — эмиттер	$U_{кэ}$	$U_{КЭ}$
Выходная мощность биполярного транзистора		$P_{вых}$
Постоянная рассеиваемая мощность коллектора		P_K

**ТАБЛИЦА СООТВЕТСТВИЯ ОБОЗНАЧЕНИЙ ЭЛЕКТРИЧЕСКИХ ПАРАМЕТРОВ
ПОЛЕВЫХ ТРАНЗИСТОРОВ, ПРИНЯТЫХ В КНИГЕ И ПО ГОСТ 19095-73**

Электрический параметр	Буквенное обозначение	
	в книге	по ГОСТ 19095-73
Ток утечки затвора	I_z	$I_{з. \text{ ут}}$
Напряжение отсечки полевого транзистора	$U_{зо}$	$U_{зи. \text{ отс}}$
Пороговое напряжение полевого транзистора	U_0	$U_{зи. \text{ пор}}$
Сопротивление сток — исток в открытом состоянии транзистора	R_0	$R_{си. \text{ отк}}$

1. Иоффе А. Ф. Полупроводники в современной физике. М.—Л., Изд-во АН СССР, 1954. 355 с.
2. Шокли В. Теория электронных полупроводников. М., Изд-во иностр. лит., 1953. 714 с.
3. Колосов А. А., Горбунов Ю. И., Наумов Ю. Е. Полупроводниковые твердые схемы. М., «Советское радио», 1965. 501 с.
4. Даилэп У. Введение в физику полупроводников. М., Изд-во иностр. лит., 1959. 430 с.
5. Spenke E. Elektronische Halbleiter, 1956. «Electronic Semiconductors», McGraw-Hill, 1958. 394 p.
6. Шпольский Э. В. Атомная физика. Т. 1. М., «Наука», 1974. 576 с.
7. Киреев П. С. Физика полупроводников. М., «Высшая школа», 1969. 590 с.
8. СВЧ-полупроводниковые приборы и их применение. Под ред. Г. Уотсона. Пер. с англ. под ред. В. С. Эткина, М., «Мир», 1972. 662 с.
9. Ноздрев В. Ф., Сеикевич А. А. Курс статистической физики. М., «Высшая школа», 1965. 288 с.
10. Смит Р. Полупроводники. М., Изд-во иностр. лит., 1962. 467 с.
11. Миддлбрук Р. Введение в теорию транзисторов. М., Атомиздат, 1960. 304 с.
12. Транзисторы (справочник). Под ред. И. Ф. Николаевского. М., «Связь», 1969. 623 с.
13. Vagnerin L. J. Stored charge method of transistor base transit time analysis. — «Proc. IRE», 1959, № 4, p. 523—527.
14. Спиридонов Н. С. Основы теории транзисторов. Киев, «Техника», 1975. 359 с.
15. Пасынков В. В., Чиркин Л. К., Шинков А. Д. Полупроводниковые приборы. М., «Высшая школа», 1973. 398 с.
16. Полупроводниковые диоды. Параметры, методы измерений. Под ред. Н. Н. Горюнова и Ю. Р. Носова, М., «Советское радио», 1968. 304 с. с ил.
17. Шокли В., Рид В. Статистика рекомбинации дырок и электронов. — В кн.: Полупроводниковые электронные приборы. М., Изд-во иностр. лит., 1953, с. 121—140.
18. Нэвон Д., Брей Р., Фен Г. Время жизни инжектированных носителей зарядов в германии. — В кн.: Полупроводниковые электронные приборы. М., Изд-во иностр. лит., 1953, с. 171—183.
19. Иглицын М. И., Концевой Ю. А., Сидоров А. И. Время жизни неравновесных носителей заряда в германии при произвольных уровнях инжекции. — ЖТФ, 1957, № 11, с. 2461—2468.
20. Стефенсон А., Кейс Р. Измерения скорости поверхностной рекомбинации на поверхности германия. — В кн.: Электрофизические свойства германия и кремния. М., «Советское радио», 1956, с. 367—372.
21. Мак-Кельви Дж., Лонджини Р. Скорость объемной и поверхностной рекомбинации носителей зарядов, инжектированных в германии. — В кн.: Электрофизические свойства германия и кремния. М., «Советское радио», 1956, с. 350—366.
22. Диткин В. А., Кузнецов П. И. Справочник по операционному исчислению. М.—Л., Гостехиздат, 1951, 255 с.
23. Янке Е., Эмде Ф. Таблицы функций. М.—Л., Гостехиздат, 1949. 416 с.

24. Shockley W. The theory of p-n junctions in semiconductors and p-n junction transistors. — «Bell System Technical Journal», 1949, vol. 28, p. 435—480.
25. Адирович Э. И., Рябинкин Ю. С., Темко К. В. Равновесное распределение потенциала, поля и концентрации носителей заряда на вплавленных переходах. — ЖТФ, 1958, т. XXVIII, вып. 1, с. 55—66.
26. Moll J. The evolution of the theory for the voltage-current characteristic of p-n junctions. — «Proc. IRE», 1958, № 6, p. 1076—1082.
27. Пикус Г. Е. Основы теории полупроводниковых приборов. М., «Наука», 1965. 448 с.
28. Gunn J. B. On carrier accumulation and the properties of certain semiconductor junctions. — «Journ. Electron. Control», 1958, № 4, p. 17—50.
29. Давыдов А. С. Квантовая механика. М., «Наука», 1973. 704 с.
30. Давыдов П. Д. Анализ и расчет тепловых режимов полупроводниковых приборов. М., «Энергия», 1967. 144 с.
31. Sah C., Noyce R., Shockley W. Carrier generation and recombination in p-n junctions and p-n junction characteristics. — «Proc. IRE», 1957, № 9, p. 1228—1243.
32. Литвинов Р. О. Влияние поверхности на характеристики полупроводниковых приборов. Киев, «Наукова думка», 1972. 116 с.
33. Замечания относительно тока Зинера в переходах. — В кн.: Полупроводниковые электронные приборы. М., Изд-во иностр. лит., 1953, с. 111—113. Авт.: К. Мак-Эффи, Е. Райдер, В. Шокли, М. Спаркс.
34. Красилов А. В., Трутко А. Ф. Методы расчета транзисторов. М., «Энергия», 1964. 224 с.
35. Шотов А. П. Об ударной ионизации в германиевых переходах. — ЖТФ, 1958, № 3, с. 437—448.
36. Matz A. Thermal turnover in germanium p-n junctions. — «Proc. IEE», p. B, 1957, vol. 104, p. 555—564.
37. Madigan J. Thermal characteristics of silicon diodes. — «Electronic Industries», 1959, № 12, p. 23—27; 1960, № 1, p. 3—5.
38. Barnes F. The forward switching transient in semiconductor diodes at large currents. — «Proc. IRE», 1958, № 7, p. 1427—1428.
39. Стафеев В. И. Влияние сопротивления толщи полупроводника на вид вольт-амперной характеристики диода. — ЖТФ, 1958, № 8, с. 1631—1641.
40. Стафеев В. И. К вопросу о вольт-амперной характеристике диода при сверхвысоких уровнях инжекции. — «Физика твердого тела», 1959, № 6, с. 848—850.
41. Авакьянц Г. М., Мурыгин В. И., Тешабаев А. Некоторые свойства диодов с большим отношением длины базы к диффузионной длине неосновных носителей. — «Радиотехника и электроника», 1963, № 5, с. 821—829.
42. Адирович Э. И., Кузнецова Е. М. О емкости и электрическом пробое p-n переходов. — «Радиотехника и электроника», 1959, № 10, с. 1708—1717.
43. Носов Ю. Р. Переходные характеристики полупроводниковых диодов. — В кн.: Полупроводниковые приборы и их применение. М., «Советское радио», 1960, вып. 4, с. 3—38.
44. Носов Ю. Р. Полупроводниковые импульсные диоды. М., «Советское радио», 1965. 224 с.
45. Степаненко И. П. Переходные характеристики полупроводникового плоскостного диода. — «Изв. вузов. Радиотехника», 1961, № 2, с. 175—184.
46. Lax B., Neustadter S. Transient response of p-n junction. — «Journ. Appl. Phys.», 1954, № 9, p. 1148—1154.
47. К вопросу об электрической формовке точечных германиевых диодов. — «Изв. вузов. Физика», 1960, № 4, с. 98—106. Авт.: В. М. Грика, С. С. Гуткин, В. М. Матошин, М. Г. Сербуленко.
48. Лосев О. В. Детектор-генератор, детектор-усилитель. — «Телеграфия и телефония без проводов», 1922, № 14, с. 380 и 385.
49. Физические основы работы полупроводниковых СВЧ диодов. М., «Советское радио», 1965. 188 с. Авт.: С. Н. Иванов, Н. А. Пенин, Н. Е. Скворцова, Ю. Ф. Соколов.

М., «Энергия», 1965. 103 с.

51. Esaki L. New phenomenon in narrow germanium p-n junctions. — «Phys. Rev.», 1959, vol. 109 (2), p. 603—604.

52. Чжоу В. Ф. Принципы построения схем на туннельных диодах. М., «Мир», 1966. 447 с.

53. Мадоян С. Г., Тиходеев Ю. С., Трутко А. Ф. Туннельный диод. — В кн.: Полупроводниковые приборы и их применение. М., «Советское радио», 1961, вып. 7, с. 35—54.

54. Сидоров А. С. Теория и проектирование нелинейных импульсных схем на туннельных диодах. М., «Советское радио», 1971. 264 с.

55. Котт В. М., Гаврилов Г. К., Баваров С. Ф. Туннельные диоды в вычислительной технике. М., «Советское радио», 1967. 216 с.

56. Эрли Д. Расширение слоя пространственного заряда в плоскостном триоде. — В кн.: Полупроводниковая электроника. М., Изд-во иностр. лит., 1955, с. 105—115.

57. Ebers J. J., Moll J. L. Large-signal behaviour of junction transistors. — «Proc. IRE», 1954, № 12, p. 1761—1772.

58. Федотов Я. А. Основы физики полупроводниковых приборов. М., «Советское радио», 1969. 592 с.

59. Эрли Д. Теоретический расчет плоскостных транзисторов. — В кн.: Полупроводниковая электроника. М., Изд-во иностр. лит., 1955, с. 116—149.

60. Транзисторы. Параметры, методы измерений и испытаний. Под ред. И. Г. Бергельсона. М., «Советское радио», 1968. 504 с.

61. Адирович Э. И., Колотилова В. Г. Переходные характеристики полупроводниковых триодов. — ДАН, 1955, т. 105, № 4, с. 709—714.

62. Эрглис К. Э., Степаненко И. П. Электронные усилители. Изд. 2-е. М., «Наука», 1964. 540 с.

63. Webster W. On the variation of junction transistor current amplification with emitter current. — «Proc. IRE», 1954, № 6, p. 914—920.

64. Sah Ch. T. Effect of surface recombination on p-n junction and transistor characteristics. — «IRE Trans.», 1962, ED-9, № 1, p. 94—108.

65. Das M., Boothroyd A. Determination of physical parameters of diffusion and drift transistors. — «IRE Trans.», 1961, ED-8, № 1, p. 15—30.

66. Справочник по полупроводниковым диодам, транзисторам и интегральным схемам. Под ред. Н. Н. Горюнова. М., «Энергия», 1972. 568 с.

67. Теория и расчет основных радиотехнических схем на транзисторах. М., «Связь», 1964. 455 с.

68. Ризкин А. А. Полупроводниковые усилители. М., Связьиздат, 1961. 120 с.

69. Герасимон С. М., Мигулин И. Н., Яковлев В. Н. Основы теории и расчета транзисторных схем. М., «Советское радио», 1963. 664 с. с ил.

70. Ван-Дер-Зил А. Флуктуационные явления в полупроводниках. М., Изд-во иностр. лит., 1961. 232 с.

71. De Witt D., Rossoff A. Transistor electronics. McGraw-Hill, 1957. 374 p.

72. Архипов В. К., Злобин А. И., Степаненко И. П. Экспериментальное исследование шумов германиевых триодов. — В кн.: Использование полупроводниковых приборов в узлах электронной аппаратуры. М., Госэнергоиздат, 1958, с. 68—77.

73. Куркин Ю. Л., Соколов А. А. Расчет схемы составного транзистора. — «Электричество», 1959, № 8, с. 66—71.

74. Кремниевые планарные транзисторы. Под ред. Я. А. Федотова. М., «Советское радио», 1973. 336 с. Авт.: В. Г. Колесников, В. И. Никишин, В. Ф. Сынов и др.

75. Дрейфовые полупроводниковые приборы. Сборник переводов. Гостехиздат УССР, 1960. 237 с.

76. Das M., Boothroyd A. Impurity-density distribution in the base region of drift transistors. — «IRE Trans.», 1961, ED-8, № 6, p. 15—30.

77. Widlar R. J. Design techniques for monolithic operational amplifiers. — «IEEE Journ. of Solid-State Circuits», Aug. 1969, SC-4, p. 184—188.

78. Сомохвалов М. М. Германиевые сплавные диффузионные триоды. М., Госэнергоиздат, 1962. 144 с.
79. Коновалов О. М. Полупроводниковые материалы. Изд. Харьковского госуниверситета, 1963. 212 с.
80. Гаврилов Р. А., Скворцов А. М. Технология производства полупроводниковых приборов. Л., «Энергия», 1968. 240 с.
81. Основы технологии кремниевых интегральных схем. Под ред. Р. Бургера и Р. Донована. Пер. с англ. под ред. В. Н. Мордюковича и Ф. П. Пресса. М., «Мир», 1969. 451 с.
82. Буш Г., Винклер У. Определение характеристических параметров полупроводников. М., Изд-во иностр. лит., 1959. 132 с.
83. Пресс Ф. П. Фотолитография в производстве полупроводниковых приборов. М., «Энергия», 1968. 200 с.
84. Bardeen J., Brattain W. The transistor a semiconductor triode. — «Phys. Rev.», 1948, vol. 74, № 2, p. 230—232.
85. Основы полупроводниковой электроники. М., «Советское радио», 1958. 579 с. Авт.: Лой, Эндрес, Зевелс и др.
86. Shenkel H., Statz H. Junction transistors with alpha greater than unity. — «Proc. IRE», 1956, № 3, p. 360—371.
87. Дьяконов В. П. Лавинные транзисторы и их применение в импульсных устройствах. М., «Советское радио», 1973. 208 с.
88. Управляемые полупроводниковые вентили. М., «Мир», 1967. 455 с. с ил. Авт.: Д. Джентри, Ф. Гутцидлер, Н. Голоныяк, Э. Фон Застров.
89. Гаринов С. А., Абезгауз И. Д. Полупроводниковые приборы с отрицательным сопротивлением. М., «Энергия», 1966. 240 с.
90. Кузьмин В. А., Сенаторов К. Я. Четырехслойные полупроводниковые приборы. М., «Энергия», 1967. 184 с.
91. Полупроводниковые приборы с отрицательным сопротивлением. Под ред. С. А. Гаринова. М., Госэнергоиздат, 1962. 240 с.
92. Березин А. С., Мочалкина О. Р. Планный тиристор с продольной структурой. — «Изв. вузов. Радиоэлектроника», 1969, № 3, с. 267—272.
93. Кузьмин В. А., Мочалкина О. Р. О зависимости коэффициентов усиления от тока в приборах типа $p-n-p-n$. — «Изв. вузов. Радиотехника», 1965, № 2, с. 165—170.
94. Шокли В. Униполярный полевой транзистор. — В кн.: Полупроводниковая электроника. М., Изд-во иностр. лит., 1955, с. 151—174.
95. Hofstein S., Heiman F. The silicon insulated-gate field-effect transistor. «Proc. IEEE», 1963, № 9, p. 1190—1202.
96. Малин Б. В., Сонин М. С. Параметры и свойства полевых транзисторов. М., «Энергия», 1967. 112 с.
97. Севин Л. Полевые транзисторы. М., «Советское радио», 1968. 184 с. с ил.
98. Martin A. Introduction a la theorie du tectnetron. — «Journ. de Phys. et le rad.», 1960, № 3, p. 24A—36A; № 7, p. 113A—121A.
99. Martin A. Field effect devices. — «Wireless World», 1961, № 5, p. 238—241.
100. Кобболд Р. Теория и применение полевых транзисторов. Пер. с англ. М., «Энергия», 1975. 304 с.
101. Кроуфорд Р. Схемные применения МОП-транзисторов. М., «Мир», 1970. 188 с.
102. Reddi V., Sah Ch. T. Source to drain resistance beyond pinch-off in metal-oxide-semiconductor transistors (MOST). — «IEEE Trans.», 1965, ED-12, № 3, p. 139—143.
103. Першенков В. С., Степаенко И. П., Яворовская Р. Я. Дифференциальные параметры МДП-транзистора в области насыщения. — «Изв. вузов. Радиоэлектроника», 1971, № 3, с. 337—343.
104. Schmidt A. Ein Beitrag zur Approximation der MOS-FET-Kennlinien. — «Nachrichtentechnik», 1969, vol. 19, H. 11, S. 129—140.
105. Басиладзе С. Г. Расчет параметров подсистемы памяти на МДП-транзисторах одного типа. — «Изв. вузов. Радиоэлектроника», 1970, № 1, с. 3—11.

106. Байков В. Д., Кармазинский А. Н., Немчинов В. М. Влияние потенциала подложки на вольт-амперные характеристики унитрона со структурой МДП. — «Известия вузов. Радиоэлектроника», 1969, № 5, с. 461—470.
107. Shannon J. MOS frequency soars with ion-implanted layers. — «Electronics», 1969, vol. 42, № 3, p. 96, 98—100.
108. Negro V., Cassidy M., Graveson R. A guarded insulated gate field effect electrometer. — «IEEE Trans.», 1967, NS-14, № 2, p. 135—142.
109. Степаненко И. П., Зайцев Б. Д. МДП-транзистор как элемент усилительных схем. — «Изв. вузов. Радиоэлектроника», 1969, № 4, с. 388—407.
110. Ходош Л. С. Технологические методы изготовления МДП транзисторов. — «Зарубежная электронная техника», 1972, № 1, с. 17—45.
111. Каганов И. Л. Промышленная электроника. М., «Высшая школа», 1968. 559 с.
112. Полковский И. М. Стабилизированные усилительные устройства на транзисторах. М., «Энергия», 1965. 216 с.
113. Мнгулин И. Н., Чаповский М. З. Усилительные устройства на транзисторах. Киев, «Техника», 1971. 324 с.
114. Полупроводниковые триоды и их применение. Под ред. Р. Ши. М., Госэнергоиздат, 1957. 288 с.
115. Шац С. Я. Транзисторы в импульсной технике. Л., Судпромгиз, 1963. 251 с.
116. Волков Ю. А. О добротности усилительного каскада на транзисторе. — «Изв. вузов. Радиотехника», 1962, № 5, с. 561—567.
117. Волков Ю. А. Анализ транзисторного усилительного каскада с индуктивной коррекцией. — В кн.: Полупроводниковые приборы и их применение. М., «Советское радио», 1963, вып. 10, с. 478—510.
118. Боде Г. Теория цепей и проектирование усилителей с обратной связью. М., Изд-во иностр. лит., 1948. 641 с.
119. Трохименко Я. К. Обратная связь в схемах с кристаллическими триодами. — «Радиотехника», 1956, № 9, с. 46—53.
120. Нагорный Л. Я. Анализ и расчет усилительных схем. Киев, Гостехиздат, 1963. 244 с.
121. Moulon J. Les transistrons dans les amplificateurs. Paris, Gauthier-Villars, 1956. 316 p.
122. Шамшин В. М. Параметры усилителей с обратной связью на полупроводниковых триодах. — «Электросвязь», 1961, № 3, с. 24—32.
123. Ковальский Е. Ядерная электроника. Пер. с англ. М., Атомиздат, 1972. 360 с.
124. Цитович А. П. Ядерная радиоэлектроника. М., «Наука», 1967, 536 с.
125. Анисимов В. И. Полупроводниковые усилители низкой частоты с непосредственной связью каскадов и обратной связью по постоянному току. — «Радиотехника», 1963, № 3, с. 65—71.
126. Лебедев В. И. Некоторые свойства эмиттерного повторителя. — В кн.: Транзисторная электроника в приборостроении. М., Оборонгиз, 1959, с. 71—94.
127. Файзулаев Б. Н. Эмиттерный повторитель в импульсном режиме. — «Радиотехника», 1960, № 1, с. 60—67.
128. Stampfl R. Transistor amplifiers with extremely high input impedance. — «Proc. NEC», 1955, vol. 11, p. 67—71.
129. Чистяков Н. И. Транзисторные резонансные усилители. М., «Связь», 1964. 32 с.
130. Радиотехнические схемы на транзисторах и туннельных диодах. Теория и расчет. М., «Связь», 1966. 512 с.
131. Трохименко Я. К. Радиоприемные устройства на транзисторах. Киев, «Техника», 1964. 416 с.
132. Ложников А. П., Сонин Е. К. Каскодные усилители. М., «Энергия», 1964. 128 с.
133. Судаков Ю. И. Амплитудная модуляция и автомодуляция транзисторных генераторов. М., «Энергия», 1969. 392 с.
134. Цыкин Г. С. Электронные усилители. М., Связьиздат, 1960. 487 с.

135. Ши Р. Ф. Усилители звуковой частоты на полупроводниковых триодах. М., Изд-во иностр. лит., 1957. 149 с.
136. Конев Ю. И. Полупроводниковые триоды в автоматике. М., «Советское радио», 1960. 447 с.
137. Коссов О. А. Усилители мощности на транзисторах в режиме переключения. М., «Энергия», 1964. 304 с.
138. Miles J. Metal-oxide-semiconductor transistors and their applications — «Mullard Techn. Comm.», 1966, vol. 9, № 83, p. 54—63.
139. Проектирование и применение операционных усилителей. Под ред. Дж. Грэма и др. Пер. с англ. под ред. И. Н. Теплюка. М., «Мир», 1974. 510 с.
140. Гальперин М. В., Злобин Ю. П., Павленко В. А. Транзисторные усилители постоянного тока. М., «Энергия», 1972. 272 с.
141. Богомолов В. Н., Николаенко Н. С., Федоров В. П. Преобразователь постоянного тока в переменный, основанный на использовании эффекта Холла. — ПТЭ, 1959, № 2, с. 134—135.
142. Нелинейная полупроводниковая емкость и ее применение. Под ред. И. П. Степаненко. М., изд. МИФИ, 1969. 147 с.
143. Шило В. Л. Линейные интегральные схемы в радиоэлектронной аппаратуре. М., «Советское радио», 1974. 312 с.
144. Филиппов А. Г. Усилители постоянного тока на полупроводниковых триодах. — В кн.: Использование полупроводниковых приборов в узлах электронной аппаратуры. М., Госэнергоиздат, 1958, с. 78—110.
145. Степаненко И. П. Анализ простейших усилительных каскадов на МДП-транзисторах. — В кн.: Микроэлектроника. М., «Советское радио», 1969, вып. 3, с. 159—172.
146. Степаненко И. П. Статический анализ простейшего дифференциального каскада на МДП-транзисторах. — В кн.: Микроэлектроника. М., «Советское радио», 1972, вып. 5, с. 190—200.
147. Кононов Б. Н. Симметричные триггеры на плоскостных полупроводниковых триодах. М., Госэнергоиздат, 1960. 160 с.
148. Gariano W. Transient analysis of junction transistors. — «IRE Trans.», 1959, ED-6, № 1, p. 90—100.
149. Васильев Е. К. Экспериментальное исследование ключевого режима полупроводникового триода на постоянном токе. — «Электросвязь», 1961, № 3, с. 18—23.
150. Gibbons J. Super-saturated transistor switches. — «IRE Trans.», 1961, ED-8, № 6, p. 443—448.
151. Анисимов В. И., Голубев А. П. Транзисторные модуляторы. М., «Энергия», 1964. 224 с.
152. Bright R., Kruper A. Transistor chopper for stable DC amplifiers. — «Electronics», 1955, № 5, p. 135—137.
153. Sparks J., Beaufoy R. The junction transistor as a charge controlled device. — «Proc. IRE», 1957, № 12, p. 310—327.
154. Абдюханов М. А., Берестовский Г. Н., Кузьмин В. А. О расчете процессов в полупроводниковых триодах методом заряда. — «Радиотехника и электроника», 1960, № 3, с. 450—459.
155. Moll J. L. Large-signal transient response of junction transistors. — «Proc. IRE», 1954, № 12, p. 1773—1784.
156. Филиппов А. Г. Об особенностях переходного процесса в транзисторном ключе при активно-емкостной нагрузке. — В кн.: Полупроводниковые приборы и их применение. М., «Советское радио», 1973, вып. 27, с. 217—226.
157. Ройзин Н. М., Маркович М. И. Накопление избыточных носителей зарядов в дрейфовых транзисторах. — В кн.: Полупроводниковые приборы и их применение. М., «Советское радио», 1963, вып. 10, с. 167—200.
158. Ройзин Н. М. Исследование некоторых явлений при работе транзисторов в импульсных схемах. — «Изв. вузов. Радиотехника», 1959, № 4, с. 462—476.
159. Кузьмин В. А. Режим насыщения в полупроводниковых триодах при больших сигналах. — «Изв. вузов. Радиотехника», 1959, № 5, с. 566—574.
160. Харли Р. Б. Логические схемы на транзисторах. Пер. с англ. под ред. Д. Ю. Панова. М., «Мир», 1965. 429 с.

161. Наумов Ю. Е. Интегральные логические схемы. М., «Советское радио», 1970. 432 с.
162. Ицхоки Я. С., Овчинников Н. И. Импульсные и цифровые устройства. М., «Советское радио», 1972. 592 с.
163. Фролкин В. Т. Импульсные устройства. М., «Машиностроение», 1966. 447 с.
164. Справочник по импульсной технике. Под ред. В. Н. Яковлева. Киев, «Техника», 1970. 656 с.
165. Степаненко И. П. Статический режим триггера с эмиттерной связью. — В кн.: Полупроводниковые приборы и их применение. М., «Советское радио», 1961, вып. 7, с. 321—340.
166. Bächle E. Eine Transistor-Kippschaltung und ihre Anwendungen. — «Elektronik», 1960, № 3, S. 68—72.
167. Степаненко И. П. Температурный и временной дрейф пороговых напряжений в транзисторном триггере. — В кн.: Полупроводниковые приборы и их применение. М., «Советское радио», вып. 8, 1962, с. 321—339.
168. Доронкин Е. Ф., Воскресенский В. В. Транзисторные генераторы импульсов. М., «Связь», 1965. 238 с.
169. Яковлев В. Н. Импульсные генераторы на транзисторах. Киев, «Техника», 1968. 443 с.
170. Агаханян Т. М., Фишман Л. Л. Исследование транзисторного блокинг-генератора. — «Радиотехника», 1963, № 4, с. 50—59.
171. Расчет и проектирование импульсных устройств на транзисторах. Под ред. М. Д. Штерка, М., «Советское радио», 1964. 567 с.
172. Орлов В. А. Анализ транзисторного преобразователя напряжения с емкостным накопителем. — В кн.: Полупроводниковые приборы и их применение. М., «Советское радио», 1962, вып. 7, с. 260—274.
173. Источники электропитания на полупроводниковых приборах. Проектирование и расчет. Под ред. С. Д. Додика и Е. И. Гальперина. М., «Советское радио», 1969. 448 с.
174. Журавлев А. А., Мазель К. Б. Преобразователи постоянного напряжения на транзисторах. М., «Энергия», 1964. 95 с. с ил.
175. Киселев Л. Н. Расчет преобразователей напряжения на транзисторах. — В кн.: Полупроводниковые приборы и их применение. М., «Советское радио», 1962, вып. 7, с. 275—288.
176. Софрошкин Ю. В. Переходные характеристики и устойчивость транзисторных стабилизаторов напряжения и тока. М., «Энергия», 1968. 168 с. с ил.
177. Кононов Б. Н. Стабилизаторы напряжения на полупроводниковых триодах. — «Приборостроение», 1956, № 10, с. 9—12.

А

Аккумуляция носителей 86
 Активации энергия 10
 Активная область базы 176
 Активный режим 184, 461
 Акцепторные примеси 10
 Алкатрон 288
 Анизотропия кристалла 7

Б

База 99, 173
 Барьер потенциальный 73, 93
 Блокинг-генератор ждущий 572
 Больцмановское равновесие 28, 71, 92
 Бриллюэна зоны 18

В

Валентная зона 13
 — связь 7
 Включение транзистора инверсное 174, 181, 470
 Внутренняя обратная связь по напряжению 181
 Время диффузии 150
 — дрейфа 249
 Время жизни 8, 46
 — задержки 179, 199, 218, 250, 346, 479
 — накопления 481
 — пролета 249
 — рассасывания 155, 483
 — свободного пробега 38
 — спада вершины импульса 341, 410
 — — динамического смещения 524
 Вырождения критерии 29
 Вырожденный полупроводник 20, 29

Г

Генерационно-рекомбинационные центры 21

Гистерезис триггерной характеристики 536
 Глубина затягивания 86
 — обратной связи 358
 Горячие носители 41
 Граничный заряд в базе 476, 486, 500, 580

Д

Дарлингтона схема 233, 392, 504
 Двойные соединения 11
 Дебаевская длина 68, 89
 Дембера эффект 76, 115, 145
 Демберовское поле 76, 83
 Дефекты кристаллической решетки 9, 46
 Динамическая нагрузка 395, 449, 504
 — отсечка 490, 519, 528
 Динамическое смещение 495, 512, 520, 524
 Диод Ганна 161
 — обращенный 171
 — с накоплением заряда (ДНЗ) 159
 Дифференциальная составляющая сигнала 442
 Дифференциальное сопротивление диода 119
 — — коллекторного и эмиттерного переходов 188, 192
 Диффузии время 150, 198, 249
 — носителей длина 82
 — — коэффициент 61
 — — — эквивалентный 140
 — примеси глубина 259
 — — коэффициент 258
 — уравнение 63, 81, 189
 Диффузионная длина 82
 — печь 258
 Диффузионное приближение 78, 81
 — сопротивление базы 193
 Диффузия подвижных носителей 27, 60

Диффузия примесей 257
Длина Дебая (дебаевская) 68, 89
— диффузионная 82
— свободного пробега 38
Добротность генератора пилообразного
напряжения 588
— транзистора биполярного 349
— — полевого 306
Донорные примеси 9
Дрейф носителей заряда 27, 36, 60
— характеристик временной 436, 549
— — температурный 428, 439, 543
Дробовой шум 226
Дырка 8, 15
Дырочный полупроводник 10

Е

Емкость запоминающая 526
— корректирующая 350
— ускоряющая 519, 523, 524
— форсирующая 494

Ж

Ждущий мультивибратор 562
— режим блокинг-генератора 572

З

Задержка фронта в блокинг-генераторе
577
— — — ключе 478, 498, 500
— — — транзисторе 179, 199, 218, 250
— — — триггере 530
— — — усилителе 355
Запоминающая емкость 526
Запрещенной зоны ширина 18, 26
Запрещенные зоны 12
Заряд в базе 148, 180, 204, 474
— — переходе 147, 474
— граничный 476, 486, 500, 580
— избыточный 476
— неравновесный 474
Заряда уравнение 475
Затвор 284
Затвора перекрытие 313
Зенеровский пробой 129
Зона валентная 13
— примесная 20
— проводимости 12

Зонная плавка 253
Зоны Бриллюэна 18
— запрещенные 12, 17
— разрешенные 12, 17

И

Избыточный заряд 476
Импульсное прямое сопротивление дио-
да 154
Инверсионный слой 71, 110, 127, 296,
Инверсное включение транзистора 174,
182, 470
Инверсный режим 461
Индексы Миллера 7
Индуктивная коррекция фронта 352
Инжекции коэффициент 117, 189, 196
— уровень 81, 100, 139
Инжекция 66, 98
Интеграл ошибок 84
Интегратор Миллера 606, 609
Ионизации температура 33
— энергия 10, 21
Ионизация примесных атомов 9, 32
— ударная 131
Ионная связь 7
Исток 284
Истоковый повторитель 315
Истощенный слой 91

К

Канал 68, 284, 296
Канальные транзисторы 284
Каскад 328
Каскадный диодный стабилизатор 629
Каскод 316, 401
Квазиуровни Ферми 30, 101
Ковалентная связь 7
Коллектор 173
Компенсированный полупроводник 21
Контакт омический 88, 110
Контактная разность потенциалов 95,
109, 295
Конуэлл — Вайскопфа формула 39
Концентрация носителей избыточная 50
— — собственная 30
— примесей эффективная 21
Коэффициент диффузии носителей 61
— — — эквивалентный 140

Коэффициент диффузии носителей эффективный 77
 — — примеси 258
 — инъекции 117, 189, 196
 — нелинейности пилообразного напряжения 588
 — нелинейных искажений 412
 — неоднородности базы 244
 — нестабильности 321
 — обратной связи (токораспределения) базовой цепи 318, 332
 — — — коллекторной цепи 334
 — переноса носителей 189
 — подавления синфазной составляющей 444, 457
 — поля 85, 244
 — рекомбинации 48
 — стабилизации 621
 — трансформации оптимальный 577
 — — согласующий 405, 418, 426
 — ударной ионизации 132, 272
 — усиления тока эквивалентный 335, 343

Критерий насыщения 463
 Критическая напряженность поля 40
 — температура 33
 Крутизна транзистора биполярного 221, 236, 423
 — — МДП 305, 310
 — — — по подложке 312
 — — полевого 287, 293
 — удельная 303
 — усилителя 361, 367, 456

Л

Лавинный пробой 131, 163
 Ловушка 22, 45, 52
 Лосева участок характеристики 161

М

Максвелла—Больцмана распределение 25
 Масса эффективная 14
 Метод пяти ординат 413, 420
 — трех ординат 413, 418, 425
 — Чохральского 253
 Модуляция проводимости 64
 — — (сопротивления) базы 140, 386
 — толщины базы 180

Молла—Эберса формулы 183
 Монокристалл 6, 253
 Мультивибратор ждущий 562

Н

Накопления время, постоянная времени 481
 — скорость 54, 58
 Напряжение обратное 87, 98
 — отпуская триггера 537
 — пороговое МОП транзистора 296, 303
 — прямое 87, 98
 — смещения нуля 446
 — спрямления зон (плоских зон) 70, 302
 — срабатывания триггера 537
 Напряженность поля критическая 40
 Насыщение скорости носителей 41
 — тока 42
 Насыщения критерий 463
 — режим биполярных транзисторов 174, 184, 278
 — — МДП транзисторов 305
 — степень 464
 Насыщенного транзистора потенциал 466
 Невырожденный полупроводник 23, 29
 Нейтральности условие 31
 Нелинейная обратная связь 498
 Неоднородный полупроводник 35, 71
 Неосновные носители 11
 Непрерывности уравнение 62
 Неравновесная ширина перехода 98
 Неравновесный потенциальный барьер 97
 Несимметричный переход 90, 92
 Низкочастотные (избыточные) шумы 226
 Носителей заряда дрейф 27, 36, 60
 — — рассасывание 51
 — — рассеяние 38
 Носители заряда горячие 41
 — — неосновные и основные 11

О

Обеднения режим 70, 294
 Обедненный слой 91, 301

Обогащения режим 70, 294
 Обогащенный слой 104, 106, 295
 Объемное сопротивление базы 194
 Объемные сопротивления слоев 467
 Обратная связь внутренняя 332, 364
 — — нелинейная 498
 — — параллельная, последовательная, по напряжению, по току 361
 Обращенный диод 171
 Однородный и неоднородный полупроводники 36
 Омический участок характеристики 136
 Оптимальное балластное сопротивление параллельного стабилизатора 625
 — сопротивление источника сигнала 420, 426
 Основные и неосновные носители 11
 Отрицательное сопротивление 133, 161, 166, 274, 278
 Отсечка динамическая 490, 519, 528

П

Перекрытие затвора 313
 Переходы несимметричные, симметричные и односторонние 90, 92
 — плавные и ступенчатые 89
 Плавка зонная 253
 Площадь усиления 347
 Плотность состояний эффективная 25, 26
 Поверхностные уровни (состояния) 46, 113, 126, 132, 295
 Поверхностный потенциал 68, 113, 295, 301

Повторитель истоковый 315
 Подвижность эффективная 77
 Поликристаллы 6, 45
 Полуметаллы 20, 29, 44, 105, 111, 166
 Полупроводник вырожденный 20, 29
 — дырочный 10
 — компенсированный 21
 — неоднородный 36, 71
 — собственный 8, 30
 — электронный 6, 9
 Полупроводники типа n и p 9, 10
 Пороговое напряжение МДП транзистора

— — — — — 296, 303
 Постоянная времени затвора 288
 — — канала (крутизны) 314
 — — коллекторного перехода 220
 — — коэффициента α 200
 — — — — эквивалентная 398, 478
 — — — β 217
 — — — — эквивалентная 343, 478
 — — накопления 481
 — — отсечки 490
 — — рассасывания 484
 — тепловая 122
 — кристаллической решетки 6
 Потенциал ионизации 33
 — насыщенного транзистора 466
 — поверхностный 68, 113, 295, 301
 — температурный 23
 — химический 27
 — электростатический 27
 — электрохимический 28
 Потенциальный барьер 73
 — — — — — перехода равновесный 94
 — — — — — неравновесный 97
 Предельная частота коэффициента усиления тока 218
 Престриктор 289
 Примеси акцепторные 10
 — донорные 9
 Примесная зона 20
 Примесные уровни 20
 Пробой зенеровский 129
 — поверхностный 132
 Проводимости зона 12
 Проводимость собственная 9

Р

Работа выхода 108
 Равновесие больцмановское 28, 71, 92
 Радиаторы теплоотводящие 237
 Распределение Максвелла—Больцмана 25
 — Ферми—Дирака 23
 Рассасывание носителей в полупроводнике 51
 Рассасывания время 155
 Рассеяние носителей 38
 Регенерация 517
 Режим активный 184, 461

Режим насыщения 174, 184, 278, 461
— — полевых транзисторов 286, 305
— обеднения, обогащения 70, 294
— сверхкритический 41
Рекомбинации коэффициент 48
Рекомбинация 19

С

Самовозбуждение мультивибратора 551
— преобразователя напряжения 620
— стабилизатора напряжения 647
Сверхкритический режим 41
Свободный электрон 14
Связь валентная (ковалентная) 7
Синфазная составляющая сигнала 442
Скорость генерации-рекомбинации 53, 135
— дрейфовая 37
— накопления 54, 58
— поверхностной рекомбинации 46, 48
— частиц тепловая 38
Следящая связь 391, 394, 594
Слой инверсионный 71, 110, 127, 296
— истощенный (обедненный) 91, 301
— обогащенный 104, 106, 295
Смещение динамическое 495, 512, 520, 524
Собственная концентрация 30
— проводимость 9
Собственный полупроводник 8, 30
Согласование сопротивлений 404, 411
Сопротивление базы диффузионное 193
— отрицательное 133, 161, 166, 274, 278
— тепловое 121, 235
Сопротивления слоев (объемные) 467
Степень насыщения 464
Сток 284
Ступенчатые переходы 89

Т

Текнетрон 288
Температура ионизации 33
— критическая 33
— удвоения тока 124, 126
Температурный дрейф 428, 439, 543
— потенциал 23
Тепловая постоянная времени 122

Тепловое сопротивление 121, 235
Тепловой ток 117, 123
Теплопроводящие радиаторы 237
Термисторы 435, 548
Тормозящее поле в базе 243
Туннельный эффект 69, 112, 129, 167

У

Ударная ионизация 131
Ударной ионизации коэффициент 132
Узлового напряжения формула 466
Уравнение диффузии 63
— заряда 475
— непрерывности 62
— Пуассона 63
Уровень инжекции 81, 100, 139
— Ферми 23, 35
Уровни поверхностные 46, 113, 126, 132, 295
— примесные 20
— энергии 11
Ускоряющая емкость 519, 523, 524
Условие нейтральности 31
— сильного сигнала 477
Участок Лосева 161

Ф

Ферми—Дирака распределение 23
Ферми уровень 23, 27
Фононы 8, 38, 41
Формула Конуэлл—Вайскопфа 39
— узлового напряжения 466
— Эйнштейна 61
Формулы Молла—Эберса 183
Функция ошибок 84

Х

Химический потенциал 27
Холла э. д. с. 254
— датчик 438

Ч

Чохральского метод 253

Ш

Ширина запрещенной зоны 13, 18, 26
— перехода неравновесная 98, 103
— — равновесная 96, 103, 107

Эйнштейна формула 61
 Эквивалентная схема транзистора для постоянных составляющих 181, 186, 215
 — — — — переменных составляющих 187, 220, 221
 — — — — α 398
 — постоянная времени коэффициента β 343
 Эквивалентный коэффициент α_{oe} 397
 — — β_{oe} 335
 — — диффузии 140
 Эксклюзия 86
 Экстракция 66
 Электронные оболочки атома 12
 Электронный полупроводник 9

Электростатический потенциал 27, 36
 Электрохимический потенциал 28
 Элементарная ячейка 6
 Эмиттер 99, 173
 Эмиттерное рассасывание носителей 486
 Энергия активации, ионизации 10
 Эффект Дембера 76, 115, 145
 — оттеснения тока 237
 — смыкания 215
 — туннельный 69, 112, 129, 167
 — Эрли 180
 Эффективная концентрация примесей 21
 — масса 14
 — плотность состояний 25, 26
 — подвижность 77
 Эффективное время жизни 46, 58

Предисловие к четвертому изданию	3
--	---

ТРАНЗИСТОРЫ

<i>Глава первая. Полупроводники.</i>	5
--	---

1-1. Введение	5
1-2. Структура полупроводников и типы проводимости	6
1-3. Энергетические зоны твердого тела	11
1-4. Зонная структура полупроводников	19
1-5. Законы распределения носителей в зонах полупроводника	22
1-6. Уровень Ферми	27
1-7. Концентрация носителей	30
1-8. Подвижность носителей	37
1-9. Удельная проводимость и удельное сопротивление	42
1-10. Рекомбинация носителей	44
Общие сведения (44). Равновесное состояние (47). Неравновесное состояние (50). Рекомбинация на ловушках (51). Время жизни (54). Поверхностная рекомбинация (57).	
1-11. Законы движения носителей заряда в полупроводниках	60
1-12. Объемные заряды и поля в полупроводниках	63
Диэлектрическая релаксация (63). Эффект поля (66). Неоднородные полупроводники (71). Квазинейтральность (74).	
1-13. Кинетика носителей заряда в полупроводниках.	75
Биполярная диффузия (75). Дрейф (78). Монополярная диффузия (79). Комбинированное движение (85).	

<i>Глава вторая. Полупроводниковые диоды</i>	87
--	----

2-1. Введение	87
2-2. Электронно-дырочный переход	88
Классификация <i>p-n</i> переходов (88). Структура <i>p-n</i> перехода (90). Анализ перехода в равновесном состоянии (93). Анализ перехода в неравновесном состоянии (97). Плавные <i>p-n</i> переходы (102). Односторонние <i>p-n</i> переходы (104).	
2-3. Специальные типы переходов	105
Переходы между примесными и собственными полупроводниками (105). Переходы между однотипными полупроводниками (106).	
2-4. Контакты металл-полупроводник	107
Выпрямляющие контакты (108). Невыпрямляющие (омические) контакты (110).	
2-5. Анализ идеализированного диода	113
Исходные предпосылки (113). Решение диффузионного уравнения (115). Вольт-амперная характеристика (117). Характеристические сопротивления (119). Температура перехода (121).	
2-6. Обратная характеристика реального диода	122
Тепловой ток (123). Ток термогенерации (124). Поверхностные каналы (126). Ток утечки (128). Эквивалентная схема диода при обратном смещении (128).	

2-7.	Пробой перехода Туннельный пробой (129). Лавинный пробой (131). Тепловой пробой (133).	129
2-8.	Прямая характеристика реального диода Ток рекомбинации (134). Сопротивление базы (135). Зависимость прямого напряжения от температуры (137). Работа диода при высоком уровне инжекции (138). Дрейфовая составляющая тока инжектированных носителей (140). Коэффициент инжекции (141). Модуляция сопротивления базы (141). Распределение токов в базе (144). Эквивалентная схема диода при прямом смещении (146).	134
2-9.	Переходные характеристики диода Барьерная емкость (емкость перехода) (147). Диффузионная емкость (148). Режим переключения (150). Установление прямого напряжения (151). Рассасывание избыточных носителей (154). Восстановление обратного тока (сопротивления) (157).	146
Глава третья Разновидности полупроводниковых диодов		159
3-1.	Точечные диоды	159
3-2.	Полупроводниковые стабилитроны	163
3-3.	Туннельные диоды	166
3-4.	Диоды Шоттки	171
Глава четвертая. Транзисторы		173
4-1.	Введение	173
4-2.	Основные процессы в транзисторе Инжекция и собирание неосновных носителей (176). Распределение носителей в базе (179). Модуляция толщины базы (180).	175
4-3.	Статические характеристики транзистора Формулы Молла — Эберса (181). Идеализированные статические характеристики (183). Реальные статические характеристики (184)	181
4-4.	Статические параметры транзистора Коэффициент передачи эмиттерного тока (188). Сопротивление эмиттерного перехода (192). Сопротивление коллекторного перехода (192). Коэффициент обратной связи по напряжению (193). Объемное сопротивление базы (194). Тепловой ток коллектора (195)	187
4-5.	Динамические параметры транзистора Барьерные емкости (196). Коэффициент инжекции (196). Коэффициент переноса (197). Коэффициент передачи тока (200). Диффузионные емкости (204). Постоянная времени базы (205). Инверсные параметры (206).	196
4-6.	Зависимость параметров от режима и температуры Зависимость от режима (207). Зависимость от температуры (210)	207
4-7.	Характеристики и параметры транзистора при включении с общим эмиттером Статические характеристики и параметры (212). Динамические параметры (216). Схема с общим коллектором (221)	212
4-8.	Разновидности эквивалентных схем П-образные эквивалентные схемы (221). Параметры транзистора как четырехполюсника (223). Сравнительная оценка (225)	221
4-9.	Собственные шумы транзистора Источники шумов (226). Коэффициент шума (228). Анализ коэффициента шума (230). Мощность и напряжение шума (232)	226
4-10.	Составные транзисторы	233
4-11.	Допустимая мощность и особенности мощных транзисторов . . .	235
4-12.	Дрейфовые транзисторы	238

Особенности дрейфовых транзисторов (238). Распределение носителей в базе (242). Коэффициент переноса (248). Динамические параметры (249).

4-13.	Элементы технологии транзисторов	252
	Получение и очистка полупроводников (252). Механическая и химическая обработка (254). Эпитаксия (255). Диффузия (257). Основные технологические циклы (261). Сплавная технология (262). Мезатехнология (262). Планарная технология (265). Фото-литография (266)	

Глава пятая. Разновидности транзисторов	268
--	------------

5-1.	Точечный транзистор	268
5-2.	Лавинный транзистор	270
5-3.	Тиристоры	274
	Динистор (275). Тринистор (280)	
5-4.	Униполярные (полевые) транзисторы	283
	Унитрон (283). Разновидности унитронов (288). Особенности реальных приборов (290). Эквивалентная схема (292)	
5-5.	Полевые транзисторы с изолированным затвором	293
	Структура и классификация (293). Физические процессы (294). Общий анализ (300). Характеристики и параметры в 1-м приближении (304). Характеристики и параметры во 2-м приближении (308). Влияние потенциала подложки (310). Эквивалентная схема (312). Схемы включения (315)	

УСИЛИТЕЛИ

Глава шестая. Статический режим усилительного каскада	317
--	------------

6-1.	Выбор рабочей точки	317
6-2.	Стабильность рабочей точки	320
	Общий анализ (320). Стабильность типовых схем (323)	
6-3.	Расчет каскадов по постоянному току	325
	Каскад с общей базой (325). Каскад с общим эмиттером (326). Каскад с общим коллектором (327)	

Глава седьмая. Усилители с емкостной связью	328
--	------------

7-1.	Введение	328
7-2.	Каскад в области средних частот	329
	Упрощенный анализ (329). Внутренняя обратная связь (332). Полный анализ (334)	
7-3.	Каскад в области больших времен и низких частот	337
	Влияние переходных емкостей (337). Влияние блокирующей емкости в цепи эмиттера (338). Совместное влияние емкостей (340). Коррекция искажений вершины (341)	
7-4.	Каскад в области малых времен и высоких частот	342
	Переходные характеристики (342). Частотные характеристики (346). Добротность каскада (347). Коррекция фронта (350)	
7-5.	Многокаскадные усилители	352
	Область средних частот (352). Область малых времен (353)	

Глава восьмая. Обратная связь в усилителях	356
---	------------

8-1.	Введение	356
8-2.	Общие вопросы однонаправленной обратной связи	356
	Основные соотношения (356). Классификация (361).	
8-3.	Внутренняя обратная связь	364
	Источник сигнала — генератор тока (365). Источник сигнала — генератор э. д. с. (366). Сравнительная оценка (368)	
8-4.	Обратная связь по переменному току	369
	Местная обратная связь по току (369). Местная обратная связь по напряжению (371). Общая обратная связь (374)	
8-5.	Обратная связь по постоянному току	380

Глава девятая. Эмиттерные повторители	381
9-1. Введение	381
9-2. Простой повторитель	382
Входное сопротивление (382). Выходное сопротивление (385). Коэффициент передачи (386). Динамический диапазон (390)	
9-3. Сложные повторители	391
Повторитель на составном транзисторе (392). Составной повторитель с внутренней обратной связью (393). Повторитель с динамической нагрузкой (394)	
Глава десятая. Каскад с эмиттерным входом	395
10-1. Область средних частот	395
10-2. Передача фронта импульса	398
10-3. Каскад с эмиттерной связью	399
10-4. Каскод	401
Глава одиннадцатая. Усилители с трансформаторной связью	403
11-1. Введение	403
11-2. Коэффициент трансформации	404
11-3. Область средних частот	406
Параметры каскада (406). Максимальный коэффициент усиления мощности (407)	
11-4. Область низших частот	409
Граничная частота и выбор индуктивности обмоток трансформатора (409). Искажения вершины импульса (410)	
11-5. Максимальная частота генерации транзистора	410
Глава двенадцатая. Мощные выходные каскады	412
12-1. Введение	412
12-2. Однотактные каскады класса А	413
Энергетические соотношения (413). Нелинейные искажения (417). Особенности каскада ОЭ (418)	
12-3. Двухтактные каскады класса В	420
Энергетические соотношения (420). Нелинейные искажения (424). Каскад с дополнительной симметрией (426)	
Глава тринадцатая. Усилители постоянного тока	428
13-1. Введение	428
13-2. Температурный дрейф	430
13-3. Однотактные усилители	432
13-4. Термокомпенсация усилителей	435
13-5. Усилители с модуляцией сигнала	436
Глава четырнадцатая. Дифференциальный каскад	439
14-1. Введение	439
14-2. Общие свойства	439
14-3. Усилительные параметры	442
14-4. Точностные параметры	446
14-5. Эволюция схемы и параметров	448
14-6. Операционные усилители	451
14-7. Каскад на МДП транзисторах	452

ИМПУЛЬСНЫЕ СХЕМЫ

Глава пятнадцатая. Транзисторные ключи	460
15-1. Введение	460
15-2. Статические характеристики ключа ОЭ	461
Режим отсечки (461). Режим насыщения (463)	
15-3. Транзисторные прерыватели	469
15-4. Метод заряда	474

15-5. Переходные характеристики ключа ОЭ	478
Задержка фронта (478). Положительный фронт (479). Накопление носителей (480). Рассасывание носителей (483). Отрицательный фронт (488). Ключ с форсирующей емкостью (493)	
15-6. Разновидности насыщенных ключей	496
15-7. Ненасыщенные ключи	498
Ключи с нелинейной обратной связью (498). Токовые ключи (501)	
15-8. Мощность, рассеиваемая транзистором в ключевом режиме	505
Глава шестнадцатая. Симметричный триггер	508
16-1. Введение	508
16-2. Статический режим	508
Условие запираания транзистора (509). Условие насыщения транзистора (510). Выходное напряжение и выходной ток (511). Статическая нагрузка (513)	
16-3. Схемные варианты симметричного триггера	514
Триггер с автоматическим смещением (514). Триггер без смещения (515). Триггер с непосредственными связями (516).	
16-4. Переходные процессы в режиме раздельных входов	516
Общее описание (517). Анализ фронтов (520). Максимальная рабочая частота (524)	
16-5. Переходные процессы в режиме общего входа	525
Общее описание (526). Анализ фронтов (530). Максимальная рабочая частота (532). Коллекторный запуск (533). Особенности применения дрейфовых транзисторов (535)	
Глава семнадцатая. Триггер с эмиттерной связью	536
17-1. Введение	536
17-2. Статический режим	536
Рабочий цикл (537). Входная характеристика (538). Анализ входной характеристики (540). Статический расчет (542)	
17-3. Стабильность порогов срабатывания и отпускания	543
Методика анализа (543). Анализ температурного дрейфа (544). Компенсация температурного дрейфа (548). Временной дрейф (549)	
Глава восемнадцатая. Мультивибраторы	549
18-1. Симметричные мультивибраторы	549
Рабочий цикл (549). Рабочая частота и ее стабильность (552). Регулирование частоты (555). Скважность импульсов и отрицательный фронт (556)	
18-2. Мультивибратор с разрядным триггером	559
Введение (559). Рабочий цикл (560). Рабочая частота и скважность (561)	
Глава девятнадцатая. Одновибраторы	562
19-1. Одновибратор с эмиттерной связью	562
Рабочий цикл (562). Статический расчет (563). Время выдержки и его стабильность (564). Время восстановления (565). Роль спускового импульса (567)	
19-2. Одновибратор с разрядным триггером	569
Глава двадцатая. Блокинг-генератор	571
20-1. Общее описание	571
20-2. Интервал между импульсами	572
20-3. Фронты импульсов	575
20-4. Вершина импульса	578
20-5. Выброс напряжения	584

Глава двадцать первая. Генераторы пилообразного напряжения	587
21-1. Введение	587
21-2. Общая характеристика и классификация	588
Параметры генераторов (588). Скелетная схема и режимы работы (589). Разновидности генераторов (590). Обратный ход (595). Особенности генераторов спадающего напряжения (596)	
21-3. Простейшие генераторы с интегрирующей цепочкой	598
21-4. Генераторы с параметрическим стабилизатором тока	600
21-5. Генераторы со следящей связью	601
Генераторы с повторительной связью (602). Генераторы с уси- лительной связью (606). Фантастроны (609)	

ИСТОЧНИКИ ПИТАНИЯ

Глава двадцать вторая. Преобразователи постоянного напряжения	612
22-1. Принцип действия и классификация	612
22-2. Анализ преобразователя с трансформаторной обратной связью	614
22-3. Особенности практической схемы	619
Глава двадцать третья. Стабилизаторы напряжения.	621
23-1. Введение	621
23-2. Стабилизаторы параллельного типа	623
Общие свойства (623). Диодные стабилизаторы (628). Транзистор- ные стабилизаторы (630)	
23-3. Стабилизаторы последовательного типа	632
Общие свойства (632). Однокаскадный стабилизатор (635). Много- каскадные стабилизаторы (637)	
23-4. Сравнение параллельных и последовательных стабилизаторов	640
23-5. Особенности практических схем	642
Регулировка выходного напряжения (642). Температурная ста- бильность (645). Зависимость параметров от частоты и режима (646)	
23-6. Методика расчета стабилизаторов	648
Приложение	650
Список литературы	653
Предметный указатель	660